

Dh	301	gvcgagagatcugaaacgaacagatccctcgagggaatacagatgctcgagcagagtcgctgcgtcg	360
Oy	361	agagcgacatccctcagagcccccgcggagagccgctgagccgctgcaacgagagaaatcgcgtgct	420
Dh	361	agagcgacatccctcagagcccccgcggagagccgctgagccgctgcaacgagagaaatcgcgtgctg	420
Oy	421	ttgcgtctgaaagagccacatctgtctgagcgagaaatcgcgtctgctgctgagccaaataaag	480
Dh	421	ttgcgtctgaaagagccacatctgtctgagcgagaaatcgcgtctgctgctgagccaaataaag	480
Oy	481	aataatcaaaatgatacaaaataaccccaatacccaataataatacaagaataaaccagatcatcta	540
Dh	481	aataatcaaaatgatacaaaataaccccaatacccaataataatacaagaataaaccagatcatcta	540
Oy	541	gtcaaaaatcttcgagcagcagagcaatgcacaaatgcacaaatgcacaaatgcacaaatgcacaaat	600
Dh	541	gtcaaaaatcttcgagcagcagagcaatgcacaaatgcacaaatgcacaaatgcacaaatgcacaaat	600
Oy	601	gttaaaaagctctagaagaaataatgcataatgcataatgcataatgcataatgcataatgcataatgc	660
Dh	601	gttaaaaagctctagaagaaataatgcataatgcataatgcataatgcataatgcataatgcataatgc	660
Oy	661	tcctctgagaaagatcagagagacagaaataaacaatgctgctgagagagacccgcatctgata	720
Dh	661	tcctctgagaaagatcagagagacagaaataaacaatgctgctgagagagacccgcatctgata	720
Oy	721	gctctgagacaaatataagagagagagagattcgaaatctgcacaaataacaaagctgagcagatacta	780
Dh	721	gctctgagacaaatataagagagagagagattcgaaatctgcacaaataacaaagctgagcagatacta	780
Oy	781	cccccacatataagaaataacattttataaagatcccatatgcacacagatgcacacagatgcacacag	840
Dh	781	cccccacatataagaaataacattttataaagatcccatatgcacacagatgcacacagatgcacacag	840
Oy	841	gaagcagacatagctctgagagagaaagaaatttataataacgctgagatactgataagagatgag	900
Dh	841	gaagcagacatagctctgagagagaaagaaatttataataacgctgagatactgataagagatgag	900
Oy	901	gagaaagcagccatccccaataccataccctgcacacattctgagctgagcccttcaaatgatacct	960
Dh	901	gagaaagcagccatccccaataccataccctgcacacattctgagctgagcccttcaaatgatacct	960
Oy	961	gagagctacatgaaatacaataaagaagcagagcttcccaaaagccaaacacataatcagccaaag	1020
Dh	961	gagagctacatgaaatacaataaagaagcagagcttcccaaaagccaaacacataatcagccaaag	1020
Oy	1021	gcaatgcagccatctgctctgcaagagataagatcctataagagatagccacagatctgaaacagga	1080
Dh	1021	gcaatgcagccatctgctctgcaagagataagatcctataagagatagccacagatctgaaacagga	1080
Oy	1081	aaagcattctgtattctttaaataacgctgagattatctacatcgcgtcccttcaaacccagccttaa	1140
Dh	1081	aaagcattctgtattctttaaataacgctgagattatctacatcgcgtcccttcaaacccagccttaa	1140
Oy	1141	gtgtcaaaagataagaccgggacatgctatgattctcaatctcccatctggagataatgagacttcaa	1200
Dh	1141	gtgtcaaaagataagaccgggacatgctatgattctcaatctcccatctggagataatgagacttcaa	1200
Oy	1201	gtagaagagagaaatgctgtgcacaaatactcaataaagagctctgcagatgctctgtatatactgt	1260
Dh	1201	gtagaagagagaaatgctgtgcacaaatactcaataaagagctctgcagatgctctgtatatactgt	1260
Oy	1261	gtgtgaaataatgagatgatacaaaataagagagagagagagagagagagagagagagagagagag	1320
Dh	1261	gtgtgaaataatgagatgatacaaaataagagagagagagagagagagagagagagagagagagag	1320
Oy	1321	atctccggaagatctgaatgatactctgcaaatgagataactctgcgaatacataaac	1380
Dh	1321	atctccggaagatctgaatgatactctgcaaatgagataactctgcgaatacataaac	1380
Oy	1381	tactctgttttaataatgaagcagacaaatgcttgagataatgagatttgaaacccagataatg	1440

Dh	1381	tacttvgtttttagatgaaagcagacagaatggttvggcatatvggaatttgaaccocagataatg	1440
Qy	1441	aagattcttctagatgctgtgcgcccagataagcagagacagttaatvgaccagtgctctcatatgccc	1500
Dh	1441	aagattcttctagatgctgtgcgcccagataagcagagacagttaatvgaccagtgctctcatatgccc	1500
Qy	1501	catctgaattcatcgccctgcgcacaaactctatttgaagaagacaatgattctctatgttgctg	1560
Dh	1501	catctgaattcatcgccctgcgcacaaactctatttgaagaagacaatgattctctatgttgctg	1560
Qy	1561	acattvgatctagctgtcgtctgaagtctcagatgaaagcaaaaataaatctgtaaccacggaagaa	1620
Dh	1561	acattvgatctagctgtcgtctgaagtctcagatgaaagcaaaaataaatctgtaaccacggaagaa	1620
Qy	1621	gagagaattggaatgcacaaatgcaaaactttctcaagagatgatgtcatatccacagacaagtctat	1680
Dh	1621	gagagaattggaatgcacaaatgcaaaactttctcaagagatgatgtcatatccacagacaagtctat	1680
Qy	1681	gtctctgtctctccgaaagaagcgtgtgtgcgatactcatataagtgaaactaaactctggaat	1740
Dh	1681	gtctctgtctctccgaaagaagcgtgtgtgcgatactcatataagtgaaactaaactctggaat	1740
Qy	1741	atatcatgtagagatctctctgcataatggaatataagagaacagaaagctggagagaaagcatataag	1800
Dh	1741	atatcatgtagagatctctctgcataatggaatataagagaacagaaagctggagagaaagcatataag	1800
Qy	1801	aactcttaaaacagcgcaaaagtgaaagatactaatctgaaacatctgatacgcctctcaagagactc	1860
Dh	1801	aactcttaaaacagcgcaaaagtgaaagatactaatctgaaacatctgatacgcctctcaagagactc	1860
Qy	1861	gagtgccatgagatgtaacacatgctataaatttggacttccacggaatataatgaaagataac	1920
Dh	1861	gagtgccatgagatgtaacacatgctataaatttggacttccacggaatataatgaaagataac	1920
Qy	1921	gtacaacggataagggcgacacgagaaagacagagagactgggtgttccatctacaacttg	1980
Dh	1921	gtacaacggataagggcgacacgagaaagacagagagactgggtgttccatctacaacttg	1980
Qy	1981	actatgaaatgcatctggagaggctgtgctctctggaattgatttaatatctccggaagaaagataacag	2040
Dh	1981	actatgaaatgcatctggagaggctgtgctctctggaattgatttaatatctccggaagaaagataacag	2040
Qy	2041	agtaatcccaagagagactgttaataatgagctcgagagggtttgtaggcacatacaacggaagaag	2100
Dh	2041	agtaatcccaagagagactgttaataatgagctcgagagggtttgtaggcacatacaacggaagaag	2100
Qy	2101	gaaatctgaaagaagaaaatctgaaagaagcctccaagagaaagcccaagaagttcattatgtctc	2160
Dh	2101	gaaatctgaaagaagaaaatctgaaagaagcctccaagagaaagcccaagaagttcattatgtctc	2160
Qy	2161	ctgtacatcagtgagtgtagagaaatctcaagaattctttagaaataatagtaagacagagaatg	2220
Dh	2161	ctgtacatcagtgagtgtagagaaatctcaagaattctttagaaataatagtaagacagagaatg	2220
Qy	2221	gacatctgtgcagtatgaaagagaccggacatgttgacatgattctttaaaataatagtggt	2280
Dh	2221	gacatctgtgcagtatgaaagagaccggacatgttgacatgattctttaaaataatagtggt	2280
Qy	2281	tgaaatcatatgaaatccagtggtttcaatacattctttaataaaataagaaagtatttaact	2340
Dh	2281	tgaaatcatatgaaatccagtggtttcaatacattctttaataaaataagaaagtatttaact	2340
Qy	2341	aaaaaataaaaaaataaaaaaataa 2365	
Dh	2341	aaaaaataaaaaaataaaaaaataa 2365	

```

RESULT      2
US-09-567-995-42
; Sequence 42, Application US/09567995
; Patent No. 6303756
; GENERAL INFORMATION:
; APPLICANT: Martelange, Val,rie

```

APPLICANT: De Smet, Charles
APPLICANT: Boon-Falleur, Thierry
TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
FILE REFERENCE: L0461/7054
CURRENT APPLICATION NUMBER: US/09/567,995
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/183,706
PRIOR FILING DATE: 1998-10-30
NUMBER OF SEQ ID NOS: 43
SEQ ID NO 42
LENGTH: 2365
TYPE: DNA
ORGANISM: H. sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (208)...(2151)
US-09-567-995-42

Query Match 100.0%; Score 2365; DB 4; Length 2365.
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgttaccgagctcgatcccttagtaacgcccagctgtgtctgtaaaagtggcggga 60
DB 1 ttgttaccgagctcgatcccttagtaacgcccagctgtgtctgtaaaagtggcggga 60
QY 61 taagaagcgtggcggggggctagcctctgtcgggctccttaagtaagcgtgtgac 120
DB 61 taagaagcgtggcggggggctagcctctgtcgggctccttaagtaagcgtgtgac 120
QY 121 ttccctgtcagctagctcttaagacgttcacggtcaggtgtgtcagaagcttgacggcaag 180
DB 121 ttccctgtcagctagctcttaagacgttcacggtcaggtgtgtcagaagcttgacggcaag 180
QY 181 acgtcggacgccccctctctgtgacaatgttcccaacgagggagctcccaagcctct 240
DB 181 acgtcggacgccccctctctgtgacaatgttcccaacgagggagctcccaagcctct 240
QY 241 acgttgggtcgttgcctagctcggcgaagctcgacagtgctcccgagcccaaggaggccg 300
DB 241 acgttgggtcgttgcctagctcggcgaagctcgacagtgctcccgagcccaaggaggccg 300
QY 301 gcggaagatctgaatcgaacacggtccctgaaggatatagtgtcggagaggtgtgcgtgag 360
DB 301 gcggaagatctgaatcgaacacggtccctgaaggatatagtgtcggagaggtgtgcgtgag 360
QY 361 agagggcaccttaggcccccggaggccgtgtgcccgtgtgtcacgaggaactgtccgctgtgt 420
DB 361 agagggcaccttaggcccccggaggccgtgtgcccgtgtgtcacgaggaactgtccgctgtgt 420
QY 421 ttgtccttgaagagccaactgtgtgtgcggtatcgtgtcgtgtgtgtgtcaaaaataag 480
DB 421 ttgtccttgaagagccaactgtgtgtgcggtatcgtgtcgtgtgtgtgtcaaaaataag 480
QY 481 aatatacaaaagtatacaacaacaacacacatcccaataatacaagaacaacagaaatca 540
DB 481 aatatacaaaagtatacaacaacaacacacatcccaataatacaagaacaacagaaatca 540
QY 541 gtcaaaatctttgtgcagcaagcaatgtcaacgaaagcaaaagcgtatatagaacaattt 600
DB 541 gtcaaaatctttgtgcagcaagcaatgtcaacgaaagcaaaagcgtatatagaacaattt 600
QY 601 gttcaaaagctagaagaataattcaattcagaatgcggaattgtatactgtcatccaaact 660
DB 601 gttcaaaagctagaagaataattcaattcagaatgcggaattgtatactgtcatccaaact 660
QY 661 tctgttggaaaaagatggaaagcacagataacaatgtgtgtcagagagatcggccattgata 720
DB 661 tctgttggaaaaagatggaaagcacagataacaatgtgtgtcagagagatcggccattgata 720
QY 721 gattgggatacaaattagaggggaaggtttgaaatgcaaaaaaagaatgggcagattca 780
DB 721 gattgggatacaaattagaggggaaggtttgaaatgcaaaaaaagaatgggcagattca 780

DB 721 gattgggatacaaattagaggggaaggtttgaaatgcaaaaaaagaatgggcagattca 780
QY 781 ccaccaatitaagaaaaactttataagaagtcacctgcacaaagtgcattgtcaaaagta 840
DB 781 ccaccaatitaagaaaaactttataagaagtcacctgcacaaagtgcattgtcaaaagta 840
QY 841 gaagcagaatgttgagaaagaataatttaataacgttggatgacttgaagatgg 900
DB 841 gaagcagaatgttgagaaagaataatttaataacgttggatgacttgaagatgg 900
QY 901 gagaacgacctatccccaatctctacgtcgaatatttgaagcgttcaatgtatacct 960
DB 901 gagaacgacctatccccaatctctacgtcgaatatttgaagcgttcaatgtatacct 960
QY 961 gaggttatgaaaaacataaaaaaggcaggtttccaagaagccaacactatccaatgcacag 1020
DB 961 gaggttatgaaaaacataaaaaaggcaggtttccaagaagccaacactatccaatgcacag 1020
QY 1021 gcatggccccattgtgttgcagaagaatataatcttaagagtagccaagctgaaacagga 1080
DB 1021 gcatggccccattgtgttgcagaagaatataatcttaagagtagccaagctgaaacagga 1080
QY 1081 aagacattgttataatgaatgccttgatltcaatcctgttccctcaaccagccttaaa 1140
DB 1081 aagacattgttataatgaatgccttgatltcaatcctgttccctcaaccagccttaaa 1140
QY 1141 ggtcaaaagaaatgaacccggcagtgtaagttaactccactcgggaatttaacacttaa 1200
DB 1141 ggtcaaaagaaatgaacccggcagtgtaagttaactccactcgggaatttaacacttaa 1200
QY 1201 gtaagaaggagaatgttgcacaatltcaataaaaggcttcggaggtgttctgttataatgt 1260
DB 1201 gtaagaaggagaatgttgcacaatltcaataaaaggcttcggaggtgttctgttataatgt 1260
QY 1261 ggttgaataatgaatgtgaacaataatagaagaagcttcaaaaaaggttataatltcaatitca 1320
DB 1261 ggttgaataatgaatgtgaacaataatagaagaagcttcaaaaaaggttataatltcaatitca 1320
QY 1321 actccggaaagatvgaatgtatctgtcaaatvgtaactctgtlcaacttgaagaataaaac 1380
DB 1321 actccggaaagatvgaatgtatctgtcaaatvgtaactctgtlcaacttgaagaataaaac 1380
QY 1381 taacttggatttagatgaagcagaacaagatgttgcataatgttttaaacccagataatgt 1440
DB 1381 taacttggatttagatgaagcagaacaagatgttgcataatgttttaaacccagataatgt 1440
QY 1441 aagaatttgttagatgtgcgccagaataggaacagatltatggccagtgctacatgtgct 1500
DB 1441 aagaatttgttagatgtgcgccagaataggaacagatltatggccagtgctacatgtgct 1500
QY 1501 catcaagtatactgcctcgcgcacaaatctatttgaagaagaacaaatgttttctatttgt 1560
DB 1501 catcaagtatactgcctcgcgcacaaatctatttgaagaagaacaaatgttttctatttgt 1560
QY 1561 acatttgatctagtgtcgttgaaglttaagtgaagcaataataatgttlaaccacaggaag 1620
DB 1561 acatttgatctagtgtcgttgaaglttaagtgaagcaataataatgttlaaccacaggaag 1620
QY 1621 gagaatgtgaatcagatgaacaactttctacaagatgtgtcatccacgaagaatgtt 1680
DB 1621 gagaatgtgaatcagatgaacaactttctacaagatgtgtgtcatccacgaagaatgtt 1680
QY 1681 gtcttcgtttctcgaaaagcgtgttcggaatcaactatacaagtgaacttaacttggaaat 1740
DB 1681 gtcttcgtttctcgaaaagcgtgttcggaatcaactatacaagtgaacttaacttggaaat 1740
QY 1741 atatacgttagatgtctctgtcatgtagaatagaacaaggaatcgggaagaacatctagag 1800
DB 1741 atatacgttagatgtctctgtcatgtagaatagaacaaggaatcgggaagaacatctagag 1800
QY 1801 aactttaaacaggaagtgtgaataactatgtcaacttgaacttagcctctagaggactt 1860
DB 1801 aactttaaacaggaagtgtgaataactatgtcaacttgaacttagcctctagaggactt 1860

```
OY 1861 gatgtccatgacgtttacacatgctctataatttctactttccacggagatattgaagaatac 1920
|||||
DB 1861 gatgtccatgacgtttacacatgctctataatttctactttccacggagatattgaagaatac 1920
OY 1921 gtaacccgaataaggcgacacggaaagacagagagactgtgtcttccattcaactcttg 1980
|||||
DB 1921 gtaacccgaataaggcgacacggaaagacagagagactgtgtcttccattcaactcttg 1980
OY 1981 actgaagaatgattggaggggttgcctctgaatttgatttaattctctggaagaagaatacag 2040
|||||
DB 1981 actgaagaatgattggaggggttgcctctgaatttgatttaattctctggaagaagaatacag 2040
OY 2041 agtattccagaaggagactctglatcacaatggcctgagaggtctgaggcacatcaacggaaag 2100
|||||
DB 2041 agtattccagaaggagactctglatcacaatggcctgagaggtctgaggcacatcaacggaaag 2100
OY 2101 gaataaggaaagaaatggaaagacctcaaggaaagcccaagaagtttcatatgctt 2160
|||||
DB 2101 gaataaggaaagaaatggaaagacctcaaggaaagcccaagaagtttcatatgctt 2160
OY 2161 cctgtactagttgggttagagaatccaagaatttttagaataatagtagaagacagatctg 2220
|||||
DB 2161 cctgtactagttgggttagagaatccaagaatttttagaataatagtagaagacagatctg 2220
OY 2221 gacatgttggcagatagaagagacccgagctgattgctatctcttaataataatagttct 2280
|||||
DB 2221 gacatgttggcagatagaagagacccgagctgattgctatctcttaataataatagttct 2280
OY 2281 tgaataatatagaatccagttgttctacttctcttaataaataatagaattacttaact 2340
|||||
DB 2281 tgaataatatagaatccagttgttctacttctcttaataaataatagaattacttaact 2340
OY 2341 aaaaaaataaaaaaataaaaaaataa 2365
|||||
DB 2341 aaaaaaataaaaaaataaaaaaataa 2365
```

```
RESULT 3
US-09-183-706-1
; Sequence 1, Application US/09183706
; Patent No. 6245525
; GENERAL INFORMATION:
; APPLICANT: Martelange, Val'rie
; APPLICANT: Boon-Falieur, Thierry
; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
; FILE REFERENCE: 10461/7054
; CURRENT APPLICATION NUMBER: US/09/183,706
; CURRENT FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: 09/122,989
; EARLIER FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 43
; SEQ ID NO 1
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-183-706-1
```

```
Query Match 13.0%; Score 307.8; DB 4; Length 323;
Best Local Similarity 98.8%; Pred. No. 2.7e-69;
Matches 321; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

OY 727 gatcaaatatgagaggaaggttgaatatgcaaaaacaaagtgggcagattaccacca 786
|||||
DB 1 gatcaaatatgagaggaaggttgaatatgcaaaaacaaagtgggcagattaccacca 60

OY 787 attagaagaaactttataaagagttccactgcccaagttgcatgtccaaagtataagca 846
|||||
DB 61 attagaagaaactttataaagagttccactgcccaagttgcatgtccaaagtataagca 120

OY 847 gatagttggaggaagaaattttaataacgttggaattgaagaatggagatggagagaa 906
```

```
|||||
DB 121 gatagttggaggaagaaattttaataacgttggaattgaagaatggagatggagagaa 180
OY 907 cgaactatcccaatccactctacactgacatttgatgacgccttcaatgttalcctgaggt 966
|||||
DB 181 cgaactatcc--aatctacctgacatttgatgacgccttcaatgttalcctgaggt 238
OY 967 atggaatacatataaagcaggtttcaaaagccacacactatcagttcacagcagtg 1026
|||||
DB 239 atggaatacatataaagcaggtttcaaaagccacacactatcagttcacagcagtg 298
OY 1027 cccatgtgtctgcaaggaatagatc 1051
|||||
DB 299 cccatgtgtctgcaaggaatagatc 323
```

```
RESULT 4
US-09-567-995-1
; Sequence 1, Application US/09567995
; Patent No. 6303756
; GENERAL INFORMATION:
; APPLICANT: Martelange, Val'rie
; APPLICANT: De Smelt, Charles
; APPLICANT: Boon-Falieur, Thierry
; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
; FILE REFERENCE: 10461/7054
; CURRENT APPLICATION NUMBER: US/09/567,995
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/183,706
; PRIOR FILING DATE: 1998-10-30
; NUMBER OF SEQ ID NOS: 43
; SEQ ID NO 1
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-567-995-1
```

```
Query Match 13.0%; Score 307.8; DB 4; Length 323;
Best Local Similarity 98.8%; Pred. No. 2.7e-69;
Matches 321; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

OY 727 gatcaaatatgagaggaaggttgaatatgcaaaaacaaagtgggcagattaccacca 786
|||||
DB 1 gatcaaatatgagaggaaggttgaatatgcaaaaacaaagtgggcagattaccacca 60

OY 787 attagaagaaactttataaagagttccactgcccaagttgcatgtccaaagtataagca 846
|||||
DB 61 attagaagaaactttataaagagttccactgcccaagttgcatgtccaaagtataagca 120

OY 847 gatagttggaggaagaaattttaataacgttggaattgaagaatggagatggagagaa 906
|||||
DB 121 gatagttggaggaagaaattttaataacgttggaattgaagaatggagatggagagaa 180

OY 907 cgaactatcccaatccactctacactgacatttgatgacgccttcaatgttalcctgaggt 966
|||||
DB 181 cgaactatcc--aatctacctgacatttgatgacgccttcaatgttalcctgaggt 238

OY 967 atggaatacatataaagcaggtttcaaaagccacacactatcagttcacagcagtg 1026
|||||
DB 239 atggaatacatataaagcaggtttcaaaagccacacactatcagttcacagcagtg 298

OY 1027 cccatgtgtctgcaaggaatagatc 1051
|||||
DB 299 cccatgtgtctgcaaggaatagatc 323

RESULT 5
US-09-208-742-3
; Sequence 3, Application US/09208742
; Patent No. 6174679
; GENERAL INFORMATION:
; APPLICANT: Kaufmann, Joerg
```

```

; TITLE OF INVENTION: CIFI50/HTAFI150 is Necessary for Cell
; TITLE OF INVENTION: Cycle Progression
; FILE REFERENCE: 1453 002
; CURRENT APPLICATION NUMBER: US/09/208,742
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO: 3
; LENGTH: 3825
; TYPE: DNA
; ORGANISM: human
US-09-208-742-3

```

```

Query Match      7.9%; Score 186.4; DB 4; Length 3825;
Best Local Similarity 49.3%; Pred. No. 6; 3e-38;
Matches 546; Conservative 0; Mismatches 556; Indels 6; Gaps 2;

```

```

Oy 965 ttatggaacacataaaaggcaggttttcaaaagccacacccatttcagtcacagcat 1024
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 926 ttatgcaccagatccgaataatctgaatacacagccacccattcaacagtcgcaagggtg 985

Oy 1025 ggcaccattgtgtcgaaggaaatagatctttagaggtagcccaactggaacagaaaga 1084
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 986 tgcctgtgcatlaagtgtgtagacatgattgtgtatgtcacaacaggtatgtggaaaa 1045

Oy 1085 catgtgtatttaattgctgagatttatcatctgtgctctcaacccagccttaaggctc 1144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1046 ctgcagccttcattctggcccatgtgttcatcttaataatgacagagaaggttgaa---c 1102

Oy 1145 aaaggaatagaccggcagctgttctaaatcccatcctcggaattgacattcaagtag 1204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1103 caggttatgtgaccaatctgcagtgattgtgtctctaccagggagccttgcacagatcc 1162

Oy 1205 aaggaaatgttgcacaatcatcatat---aaaggcttcggaagcttgtgtataggtg 1261
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1163 atccagaatgttaagcgggttttggaaaagcatataatcttcacatcggtgcgtataggg 1222

Oy 1262 gtggaataagatgaacaaatgaagaagcttaaaaaggtgtatatacataatgtgcaa 1321
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1223 gaggagatgtctgggagcagggccacagcccttcagaggggcagagatgtgtgtgta 1282

Oy 1322 ctcccggaagatgtaatgctgcaaatgtgaactcgtcaatcttgaagaataataacct 1381
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1283 ccccaaggtcagcatgataatcatgtgaaaaaagaagcaccacatcttcaaaaggtctctt 1342

Oy 1382 acttggttttagatgaagacagacaagtgttggacaatgggattggaaccacagataatga 1441
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1343 acctgtgttttgaataagacagatcgaaatgtttgaactgtgatttgaatcaccaagttcgat 1402

Oy 1442 agacttgttgaatgtcggccagatagacagatgaatgacaggtgtactacatgtgcctc 1501
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1403 ccatagcaagtcagtgtctgcctgacagagagatctctcttaattagtgcaacttttcgga 1462

Oy 1502 attcagttcatcgcctcgcacaaatcttaatttgaagaaccaaataatgttctatgttggta 1561
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1463 agaaagatgaaagtgtgcagagacatccctgacacccattcgaatgtgtgcagggag 1522

Oy 1562 catgtgattcagttgtctgtaagtctgaagcaaaataataatgttaacacacaggaag 1621
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1523 atattggagagcaaatgaagatgtgacacagatgttgagatcttcattctcgagacca 1582

Oy 1622 agaaatgagatcacatgcacaaactttctacagatgtgtcatccacagacaagaatcatgt 1681
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1583 gtaaaatgaaactgtctacccggcgtctgttgaatttacccttcagggaagtgtctctcc 1642

Oy 1682 tcttcttctcgaaaaactgtctgcgataacttacaagtgtacataatcttggaaata 1741
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1643 tcttgttactaaaaaagcacaatgtcgaagagtagcgaataacttaaacagaggagtc 1702

Oy 1742 tatcagtagagtcctgtatgagatagaaagagagatctgggagagaagatcttgaga 1801
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1703 ataacttggctgtctcatctggtgatatgtatcagtgtagagaacaagttcatcttcag 1762

```

```

Oy 1802 actttaaacacggcaaaagtgaagataactaattgcaactgtatctagcctcttagagacttg 1861
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1763 acttaagaaaaagacatccacgctccgtgggacacaaagtgttgaagccgtgtgtcgg 1822

Oy 1862 atgtccatgacgttacacatgtctataatcttgcatttccacggaatattgaagaatcg 1921
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1823 acattccttaattgaactgtcatctaactatgtatgtgcagcggacattgtatcgaca 1882

Oy 1922 tacaccgaatagggcgacaggaagagcaggaagactgtgttccattaaactttga 1981
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1883 ctcaagatattgcccgaacaagaagagcgggttgaaaggtgtgcctctataccctacca 1942

Oy 1982 ctgaagaatgtattgaggtgttcctctgaattgaatataatctcggaaagaacaaatcga 2041
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1943 ctcccaagagcagaatattgtctgtgtgcctgcgcgaacttggaaagagcacaatcac 2002

Oy 2042 gttatccagagaggcgtgtatcatatggc 2069
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2003 acgtttcagaagacctcctagatctggc 2030

RESULT 6
PCT-US96-05320A-894
; Sequence 894, Application PC/TUS9605320A
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences
; APPLICANT: 9410 Key West Avenue
; APPLICANT: Rockville, MD 20850
; APPLICANT: United States of America
; APPLICANT: Johns Hopkins University
; APPLICANT: 720 Rutland Avenue
; APPLICANT: Baltimore, MD 21205
; APPLICANT: United States of America
; APPLICANT: Mark D. Adams
; APPLICANT: Owen White
; APPLICANT: Hamilton O. Smith
; APPLICANT: J. Craig Venter
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genom
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20003-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05320A
; FILING DATE: April 12, 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: June 7, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Eric K. Steffe
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488, 014PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SRO ID NO.: 894:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1254 base pairs
; TYPE: nucleic acid

```

STRANDEDNESS: double
TOPOLOGY: linear
PCT-US96-05320A-894

Query Match 6.48; Score 151; DB 5; Length 1254;

Best Local Similarity 49.68; Pred. No. 3.9e-29;

Matches 495; Conservative 0; Mismatches 495; Indels 7; Gaps 4;

```
OY 981 aaaggcaggtttccaagaagccaaccattatcagtcacaaggcattggtctgttgcga 1040
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84 AAAAAAAGCTTTGATTTTGTATCCCAATTCAGCCTTATCCTTGCGCTATCGTTTAA 143
OY 1041 aggaatagacctatagagtagccagactggaacaggaagacattgttattaat 1100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 144 TGGACGAGATGTCCAGACACACTCAAACTGTACAGGCAAGCAATGCGCTTTTAAAC 203
OY 1101 ggcgtgattcatcattgttccctcaaccagacctaaagtcagaagatagaccgg 1160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 GCGTACTTTTCAACCATCTTTTAACTCACCAGATCCTTAATCTTAATCCTCACCAG 263
OY 1161 catgttagtcttaacctccacctcggaattgacctcaagt--agaaggagaatgtg 1217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 AGCTTTGATTTTACGACCTTACTCGAATAATTACGGTACAGATTAGTAAAGCAGAA 323
OY 1218 caaatatcatataaaggagctcgaggtgttgtgtatagtggttgaatatagatga 1277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 324 TCTTGCAAAAGCGAGTGGATTAAAGACGCGCACTTGCCATGCTGCGATGCTTATGAT 383
OY 1278 acaaatagaagagcttaaaaaagtgtagatatactaatgtcaacctccggaattga 1337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 384 ACAACTCAACGAGATTGAGCGTGGCGTCGATATTTTGTGATACGACGGGGGAGTCAT 443
OY 1338 tgatctgcaaaatgagtaactctgcacatctgaaagatlaaacctacttggtttatga 1397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 444 TCATTATGTGAACAAGCGCTAATTGTTTACATCAATCCAAAGTTCGTGCTAGATGA 503
OY 1398 agcagacaagaagtgtagacatgtagatttgaacccagatlaagaaatttctgttag 1457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 504 ACGAGATCGAATGTTTATCTTGCGTTT--ATCCGTGATTAATCTTATTTATGCGTAAA 561
OY 1458 ggcgcagatagagaacagltatgaccagtgtaacatgagcctatcagattcagct 1517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 562 TGCCTCGCTCCGCAAGCTCGTTAAGCATGTTATTTTACGAGACGCTTCTTATAAG 621
OY 1518 ggcacacattatcttgaagaacaaatgattgtcatgtgtgtcatcttgatctgtgc 1577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 622 CGTGATTTAGCATTTGAGATGATGAAATGAACTGAAATATTTGAAATTCACAGACAA 681
OY 1578 tgta-agltcagtgaaagcaaatataatgttaaccacaggaagaaatgagatcaca 1636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 682 AAACACAGGACACCGAATTTAAGAAAGAACTTTTATCCATCTATCAGATAAATGGCA 741
OY 1637 tgcacactctctcagagatgt-cattcacagacaaagcatgtgtctgtcttcga 1695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 742 CTTCCTTAACCTTAATGCAAGATGAAATGCTTAAGCGCTGTATGTATTTGGAAATAC 801
OY 1696 aaagcgtgtcggaatcactatcaagtgacctaacttggaaatataatcagtagact 1755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 802 AAACATTCGTTGTGAAGAAATTTGGGCTATTGTGGCGCTGATGGACATCGTGCGTT 861
OY 1756 ctgcattgagatagagaacagagatgtaggaagaagatagagaactttaaacaggc 1815
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 862 CTGACTGGCGATGTAGCACAAGAAAAACCTTTATCGTTATTTAAACAAATTTACGATG 921
OY 1816 aaagtgaataactaatgtcaactgtatagcctctagaggaactgtgtcattcagct 1875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 922 GATTTGATATTTTATGAGCAACAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 981
OY 1876 acacatgctataatctttagacttccaaggaatatagaagaatcgaacacgaataggg 1935
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 982 ACGCATGTTTCAATTAATGATTTACCCGATGATCAGAAAGATTAATTTCCACGAA 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
OY 1936 cgcacggaaagacagggagactggtgttcccatca 1972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1042 CGTACTGAGACGACGCGGAAGTGTGTTCCATTA 1078
```

RESULT 7

US-09-058-489-90

; Sequence 90; Application US/09058489

; Patent No. 6103886

; GENERAL INFORMATION:

; APPLICANT: Whitehead Institute for Biomedical Research

; APPLICANT: Lahn, Bruce

; TITLE OF INVENTION: Genes In the No. 6103886-Recombining Region of

; FILE OF INVENTION: the Y Chromosome

; FILE REFERENCE: WH197-08PA

; CURRENT APPLICATION NUMBER: US/09/058,489

; EARLIER FILING DATE: 1998-04-10

; EARLIER APPLICATION NUMBER: 60/041,877

; NUMBER OF SEQ ID NOS: 91

; SOFTWARE: FastSeq For Windows Version 3.0

; SEQ ID NO 90

; LENGTH: 2319

; TYPE: DNA

; ORGANISM: Human

US-09-058-489-90

Query Match

Best Local Similarity

Matches 459; Conservative

0; Mismatches 446; Indels

18; Gaps

3;

6.18; Score 143.4; DB 3; Length 2319;

49.78; Pred. No. 4.4e-27;

Matches 459; Conservative

0; Mismatches 446; Indels

18; Gaps

3;

1165 ttggttctaactccaccctcggaattagcaactcaagttagaaggaagatgttcaaatat 1224

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 873 ttggtttgcccccaacgaagaattgtcgtatcagatctataggaagccaagaatttc 932

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 1225 tcatataa---ggctccggaagtgtgtgtatagtggtgtggaataagatgaacaa 1281

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 933 tcttaaccgacttagagtttgccttctgtgtattatagtggtgtcgtatattgtgcagcag 992

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 1282 atagaagagcttaaaaaagtgtagatatactaatgtcaactcccggaagattgaatgat 1341

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 993 attcggagacttagaacgtagatgtcactgttagtagcaccatccagagcgtctagtgat 1052

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 1342 ctgcacaaatgagtaacttgttcaacttgaagaataataactactgtgttttagaagaac 1401

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 1053 atgattgaaagaggaagaatctggaatctgcaagtaacttagtgttgaatgaagct 1112

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 1402 gacaaagatcttgaaatcttgaaacccagataatgaagaatttcttagatg---- 1456

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 1113 gatagatgtcgtatagatgttgaactctcagatacgtctgatatagtgtgaacaagatact 1172

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 1457 -----tgcgccagatagagcagacgttatgaacagtgtaacttgcctcattcagtt 1509

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 1173 atgcacacaaaggagcgtctgcacacacaaatgatttagtgtaacttccctaagaataa 1232

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 1510 catcgcctcgcaaaatcttatttgaagaaccaaagtatgtctatgtgtgtacattgabt 1569

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 1233 cagatgtcgtctgtgtaacttlttgatcatatacttlttgcgttaggaagagtaggt 1292

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 1570 ctatgttcgttagatctcagatgaagcaaaatataatgttaaccaacaggaggaagaaatgg 1629

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 1293 tctactcttgaga---acatcaacagaagaatgttcttggttggaagacttagataaacgg 1349

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 1630 agtacaatgcaaatcttctacaagatgtlcatccaagaaagaatgattgtcttcgtt 1689

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 1350 tcaattctactgtgacattttagtgcaacaggaggtatcacttaacttagtgtttgtg 1409

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 1690 tctcgaagaacgtgtgtcggtgaacttcaaggaacttaattcttggaataatctcagta 1749

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 1410 gagacaaaagaagagagacttcccttggaagattcttatactacgaagatagatgtctgt 1469

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Query Match	6.1%	Score 143.4	DB 3	Length 4416
Best Local Similarity	49.7%	Pred. No. 5.9e-27		
Matches 459	Conservative	0	Mismatches 446	Indels 18
			Gaps	3
Qy	1165	tttgtcttaactcccaatccggaaatagacttcaagttagaagagaaatvgtgcaaat	1224	
Db	873	ttgttttagccccaacaagaagattgtgttacaagatcatatgaaagcagaataatt	932	
Qy	1225	tcataataa---gggcctcggaagtgtgtgtatagatgtgtggaatagagaagaa	1281	
Db	933	tcctaacgactctagatgctgcctctgtgtagtattagttgtgtgtcgtatattgtgtcagag	992	
Qy	1282	atagaagagcttaaaaaaggtgtagatatcatatgtcaactcccggaagtgtaatgat	1341	
Db	993	atccggaactctagaacgtggaatccactgtttatgaacacatccagaagctgtagt	1052	
Qy	1342	ctgcaaatgtgaacttcgcgaactctcggaatatcaaacacttggttttgaagaaga	1401	
Db	1053	atgattgaaagaagaagaattgattagactcttcgcgaactacttagtgtgtgaatgaagt	1112	
Qy	1402	gacaagaagtgtgacatcgggaatttgaaccacagataatgaagatttgtttatagv	1456	
Db	1113	gataagagatcgttgatattgtgatttgtaactctagaatcgtctgattagtgtgaacaaataact	1172	

```

1  RESULT 9
2  US-09-058-489-14
3  : Sequence 14, Application US/09058489
4  : Patent No. 6103886
5  : GENERAL INFORMATION:
6  : APPLICANT: Whitehead Institute for Biomedical Research
7  : APPLICANT: Lahn, Bruce
8  : APPLICANT: Page, David
9  : TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
10 : FILE REFERENCE: wh197-08pA
11 : CURRENT APPLICATION NUMBER: US/09/058,489
12 : CURRENT FILING DATE: 1998-04-10
13 : EARLIER APPLICATION NUMBER: 60/041,877
14 : EARLIER FILING DATE: 1997-04-11
15 : NUMBER OF SEQ. ID NOS: 91
16 : SOFTWARE: FastSeq for Windows Version 3.0
17 : SEQ ID NO 14
18 : LENGTH: 3408
19 : TYPE: DNA
20 : ORGANISM: Human
21 : US-09-058-489-14

```

Query Match	5.7%;	Score 135.4;	DB 3;	Length 3408;
Best Local Similarity	49.2%;	Pred. No. 5.7e-25;		
Matches 454;	Conservative	0;	Mismatches 451;	Indels 18;
				Gaps 3;

OY	1165	ttatgtctaaaccccaaccccgagatttagaactctaaagaaagaaagatgtgtgcaaat	1224
Db	1664	ttgtattatgtcccaacgagagaggtttgtgcagaaacagatctctgaagaagccaaatatt	1723
OY	1225	tcaataaa---ggactcggagtgltctgtatagtgtgtgaaatagaatgaacaa	1281
Db	1724	tcaataccgactctagagatttgcctcttcggtgttatagtgtgtgcccattatgttcagacg	1783
OY	1282	atagaagaaactaaaaaaaggtgtatagaatacaaatattgaactcccgaaagatgaaatg	1341
Db	1784	attcggagacttggaaatgtgatagtccatttgttagtagcaactccagaagcctgattg	1843
OY	1342	ctggcaaatgtgtaacttcgtccaatcgtgaagaatataacctattgtgtttaaagtgaag	1401
Db	1844	atgtatgaaagaagaaagatcttgatttagactcttgcgaatctctgtgtgtatagatgaagct	1903
OY	1402	gaacaaatgtgtgacatgaggatatttgaaaccccaagaatagaagatttgtataga-----	1454
Db	1904	gactcgaaatgtgtgataatgtgggttttgctgcctgaagatctgtagaatagtcgaacaagatact	1963
OY	1455	-----tgtgtgccagatagcgcgaagttatagaacagatgtctatagcctcaatccagtc	1509
Db	1964	atgtcctccaaagagtggtccgcgcacactaaagaatgtattgttactcttccctaagaataa	2023
OY	1510	catcgcctcgcacaaactctatttgaaagaacaaatgatactgtatgtgttgcattgagat	1569
Db	2024	caagatgtcgtcgtcgtatctcttaagatatactcttctgtgtcgttagaagagttgtgc	2083
OY	1570	ctaatgttcgtctaaagttcacgaaagaaataataattgtaaaccccgagaagaagaatgg	1629
Db	2084	tctaacctctgaaa---acatccacacgaaagatgatttgggtgtgaagaatcagaacaaagc	2140
OY	1630	agtcacatgtcaaaactttcttaacagagatgtatcatccacagaacaangtcatgtctctgt	1689
Db	2141	tcaattctgtctgtgacccctctaaatgcaaacaggtcaagatctcaatgtgacttattgttctg	2200
OY	1690	tctcgaanaaagctgtgtcgtgtacaaactatcaagtgaccataactctgaatatatcagta	1749
Db	2201	gagacccaanaaagggtgagagattctctcgtgaggaattctctataccaatgaagaatagcagatg	2260
OY	1750	gagctctcgtctgtgagatagagaacagagagatcgtggagaagaagcattagaacttaaa	1809
Db	2261	accagacatccatctgtgagaccgcttctcgaagagatagaaagaagcccttaccagatgccgc	2320
OY	1810	acaggcaaaagtgaaatataactatgtgaactctgaactccgctcgtgagacttattgccat	1869
Db	2321	tcaagaaaaagcccaattttagtgtgtctacagcagtagtcagcaagagactgtgaacttca	2380
OY	1870	gacgtttaaagctgtctataattttgaacttccacgggaatatttgaagaataagttacacga	1929
Db	2381	aattgtgaacaactgttataccaatttctgactgtgcgaagtgatattgaagaataatgtataactgt	2440
OY	1930	ataagtcgcaacggtgaagagatggagagagactgtgttctccaattcaactctgtcagaat	1989
Db	2441	attgtgtcgtaaacgagcgtgaagaacactgtgcctgtgcacccatcttttaagagaggg	2500
OY	1990	gatttggagaggtgtgcctctgaatttgtataattcttgaaagagacaatcagaattatcca	2049
Db	2501	aacataaaatatctactaagatttgttctgtgactctctctgtgtgaagctaaacaagaagtgcgc	2560
OY	2050	gaagagactgtatcaatgtgcga	2072
Db	2561	tcttgtttagaanaacatgtgctta	2583

```

? TITLE OF INVENTION: Genes in the No. 6103886-Recombining region of
? TITLE OF INVENTION: the y Chromosome
? FILE REFERENCE: WH197-08pA
? CURRENT APPLICATION NUMBER: US/09/058,489
? CURRENT FILING DATE: 1998-04-10
? EARLIER APPLICATION NUMBER: 60/041,877
? EARLIER FILING DATE: 1997-04-11
? NUMBER OF SEQ ID NOS: 91
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 13
? LENGTH: 5322
? TYPE: DNA
? ORGANISM: Human
? US-09-058-489-13

```

[illegible]

Db 2441 attgctgacgagcagcgtgtagaacccttgacctgcaacctctcttaacgagag 2500
QY 1990 gattggagggtgctcctgaattatataatctcgaaagcgaatcgaatccca 2049
Db 2501 aacataaataataacaaagattcgtgacacctctgttgaaagcaagaagtgcg 2560
QY 2050 gaaggactgtatcaatgctga 2072
Db 2561 tctggttagaaaacatgctta 2583

RESULT 11

US-09-039-773A-1

Sequence 1, Application US/09039773A

Patent No. 6100388

GENERAL INFORMATION:

APPLICANT: Casas, Ivan

APPLICANT: Jonsen, Hans

APPLICANT: Mjlsam, Bo

APPLICANT: Roos, Stefan

TITLE OF INVENTION: Lactobacilli Harboring Aggregation and Mucin

TITLE OF INVENTION: Binding Genes As Vaccine Delivery Vehicles

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESS: Standley & Gilcrest

STREET: 495 Metro Place South, Suite 210

CITY: Dublin

STATE: Ohio

COUNTRY: US

ZIP: 43017

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 Inch, 1.44Mb storage

OPERATING SYSTEM: MS-DOS Version 6.22

SOFTWARE: Microsoft Word Version 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/039,773A

FILING DATE: 16-MAR-1998

CLASSIFICATION: 536

PRIOR APPLICATION DATA: No. 6100388 applicable

ATTORNEY/AGENT INFORMATION:

NAME: Donald O. Nicky

REGISTRATION NUMBER: 29,092

REFERENCE/DOCKET NUMBER: 1229-005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (614) 792-5555

TELEFAX: (614) 792-5536

TELEX: No. 6100388 applicable

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1800 base pairs

TYPE: Nucleic acid

STRANDEDNESS: Double

TOPOLOGY: Circular

MOLECULE TYPE: Genomic DNA

DESCRIPTION: Genomic DNA sequence and deduced amino

DESCRIPTION: acid sequence of bacterial aggregation

DESCRIPTION: protein

HYPOTHETICAL: NO

ANTI-SENSE: Yes

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM: Lactobacillus reuteri sp

STRAIN: 1063

CELL TYPE: Unicellular organism

US-09-039-773A-1

Query Match 5.1%; Score 120.6; DB 3; Length 1800;
Best Local Similarity 46.6%; Pred. No. 2.5e-21;
Matches 476; Conservative 0; Mismatches 524; Indels 21; Gaps 2;

QY 971 aaaaataaaagagcaggttttcaaaagccaactatctagtaacgagatgccc 1030
Db 221 AAGCATCAAAAGAGCGGATACGAAGAGCAACCAATTCAGAAAGCAATTC 280
QY 1031 tctgttgcgaagaatagatctatagagtagtaccagacatggaacaggaaagacatgt 1090
Db 281 TGGTCTCTGAGGGTAAGAGATGTATGTGCAAGCAGACATGGAATGGTAAGAC----- 335
QY 1091 gtataataagcctggatattatcatcttgcctctcaaccagccttaagttcaagga 1150
Db 336 -----GCTGCTTTTGGGTGCCAATTTATGAAAAGCTTGATACGAAATCCCA 385
QY 1151 atagaccgagatgttagttctaaacctccaccccggaattagcactcaagtagaagag 1210
Db 386 ATATTCAAGCAATTAATCATTTTCAACACAGCGAATTAAGCATTCAGACCAAGAAAGAC 445
QY 1211 aatgttgcgaatatctcaataaaggcttggagtggttctgtatattgtgtggaata 1270
Db 446 TTTATCGCTTAGTAAAGATTAACATGTTCCGTCGAGGTATGTATGGTGGGCAATTA 505
QY 1271 gagaatgaacaataagagagcttaaaaaagtgtagaatacacaattgcaaacctccgga 1330
Db 506 TTGCGCGCAAAATTAAGACCTTGAAACCAACCCCAAAATTCGTGGGGACCCCTGGAC 565
QY 1331 gattgaatgactcgcgaatgagtaacttcgcgaatcggaagaatataactactgtgtt 1390
Db 566 GGTATCGTAGCAATTTAAACCGTCATACAGTTAACTTGACCACTTAAGACCTGTGTC 625
QY 1391 tagatgaagcagaagaatgttgcgaatcagagattgaaacccaataatgaagaattgt 1450
Db 626 TCGATGAAGCAGATGAATGCTAAACATGGGATTCCTTAGAAGATTAATGAATCATCATCA 685
QY 1451 tagatgtgcgccagaatagcagacatgagacagtgctacatgacctcaatcagctc 1510
Db 686 AGGAACACCAAGATGATCGCAAACTTGTCTTCAAGCAACATCCACCAAGAAATCA 745
QY 1511 atgcctcgcgaacaatcttatctgaagaacccaatgattgtctatgtgttaccttgatc 1570
Db 746 AGCAATTTGGGTTCAATTTATGCTGATCCGGAACCTTGCGCATCAAGCGCAAGGAT 805
QY 1571 tagttgtcgaagtctcagtgaaagcaaaataatctgtaaccacgcggaagaagaatga 1630
Db 806 TGACTACTGACTTATGTTGATCAGTACTA-----TGTTCCGCTCCTGACTATGAAGAT 859
QY 1631 gtcaatgcaaaactttctcaagagtagtcaatccacagaacaaatcattgtctcgttt 1690
Db 860 TTGACATCATGACCGCTTAATTTGATGTTCAAGATCCTGACTTAACAAATTTGCTTGCTC 919
QY 1691 ctgaaaagcctgttcgcgaatcaatataaagtgaacttaacttggaatatatacagtag 1750
Db 920 GGACAAAGCGCGGTATGATGAATTTGTGAAGCGCTTGATTTGGCGCTGCTCAATGACG 979
QY 1751 agtctcgtcgtgagtagtaagaacagagagctcggaagaacatagaagaacttataa 1810
Db 980 CTGGTATCATGATGTGACTTACTACAGATTAAGCCCTTCTTAAGATCATGTGGAATTAAGA 1039
QY 1811 caggcaagtagaataactaactgaactgaactgaactcctctagagagactgtgtcctatg 1870
Db 1040 ACAATGAACCTTGATATCTTAATGTTGCAACAGATGTGCGCGGCTTAAGCAATTCGCG 1099
QY 1871 acgttacaatgctataattgtgacttcaaggaatataagaataatgtagtaacgga 1930
Db 1100 GGGTTACGATGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1159
QY 1931 taaggcagacggaagaagcaggaagagactgtgttctcaatcaacttgaactgaagaatg 1990
Db 1160 TTGGCCGAAACAGCAGCGCGCATCATACGGGATCTTAACTTTGACTCCAATG 1219
QY 1991 a 1991
Db 1220 A 1220

```
RESULT 12
US-09-356-952-12
/ Sequence 12, Application US/09356952
/ Patent No. 611763
/ GENERAL INFORMATION:
/ APPLICANT: Borjack-Sjodin, Ann
/ APPLICANT: Margalit, S. M.
/ APPLICANT: Bor-Sogli, Dafna
/ APPLICANT: Cole, Philip
/ APPLICANT: Kuriyan, John
/ TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: 600-1-228N
/ CURRENT APPLICATION NUMBER: US/09/356,952
/ CURRENT FILING DATE: 1999-07-19
/ EARLIER APPLICATION NUMBER: 60/093,631
/ EARLIER FILING DATE: 1998-07-21
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: Patentn Ver. 2.0
/ SEQ ID NO 12
/ LENGTH: 43676
/ TYPE: DNA
/ ORGANISM: Saccharomyces cerevisiae
US-09-356-952-12

Query Match          4.6%; Score 107.8; DB 3; Length 43676;
Best Local Similarity 47.1%; Pred. No. 1.9e-17;
Matches 488; Conservative 0; Mismatches 522; Indels 27; Gaps 4;

QY 955 tatcctgaggtatgtaagaaacattaaaggcaggtttccaaggccaacactatccag 1014
DB 21111 tctcgctcggtcttaaggcccttgaaagtttggttaagtcagtcacgcttccctattcaa 21170
QY 1015 tcaacgagatgagcccatgtgtgtgcaagaatagattcttatagagtaagcccaagctga 1074
DB 21171 agcgccaacatcccatctgcttattgtgtaagaacacatcagcagtgctgtgctgagct 21230
QY 1075 aacaggaagacatvtgttatttaatgacctgatttcaatcctgctccccaagcagc 1134
DB 21231 tccggtaaagactgtcgttttaagatcccatlaacgcagcgcttgtgtataaac--- 21286
QY 1135 cttaaaagctcaaaagatagaccggcagtgatgtcttaactccactcgaggaattagca 1194
DB 21287 -----agccaataatcgcttccacacagattatgttctatctgcccactcgtgagct 21341
QY 1195 ctcaagtaagaagaaatgtgtcaaatatcattcataaagggttcgga-----gtgtt 1248
DB 21342 atccaaagtcgtgacgtgtgttaaacaaattgcacglttcgtctccggtataaacctt 21401
QY 1249 tgtgtatattgtgtgtaagaatagagaatgaacaataagaagcttaaaaagtgtagat 1308
DB 21402 ctgagcggtgtgtgttgaaccttgagacacacagaacaaatgltgaatcctgctcgagc 21461
QY 1309 atcataattgcaactcccggaagattgatatctgcacaatga-----gttaactgtccaat 1365
DB 21462 atcgtcattgtctacccagtgatcatctatgctataacgaagaaatcgaagatttcaat 21521
QY 1366 ctgaagaaataaactactgtgtttatagaatgaagcaagaagatgtgtgacatvggattt 1425
DB 21522 gtgagactcagtaagatctcgtgtatgagtgagcgcagatagaatgltgaagaagaagttt 21581
QY 1426 gaacccagataatgaagatttgttaagatgtgtgcccagataagacagatgattagacc 1485
DB 21582 caaagatgaactgaaacgaataattatggtcctattacaaacgaatagaacgaacattgtt 21641
QY 1486 agtgcataatgacctcaatcagttcacctcgcccgacaacatcttcttgaagaagaacatg 1545
DB 21642 tctgtacaatgaacatccaaatlaaagttaagttagtcttcttctcaaaaaaacagta 21701
QY 1546 atgtgtcattgtgttaacatgtgatactgaagtctgttaagttcaatgtgaagcaaatataat 1605
DB 21702 aggatataatgtatctctccaaagaagcgtctactaagtgtgacacagaatctgttgt 21761
```

```
QY 1606 gtaaccaccgaggaagaatgtagtcaatgcaaacctttctacagaggtatgtatcc 1665
DB 21762 atctgtataaagaagacattgaagccgctgtgtatttaatttgatagaatgtgat 21821
QY 1666 ac-----agacaagatcatgtcttctgttcttcgaagaagctgtgtgacactta 1716
DB 21822 ccaacggtcaaaagaagatgtcgtttttgtgtgtaagaagaacagctctcattagta 21881
QY 1717 tcaagtagactatactctggaatatatactagtagagttctctgcatgtagatagaagacg 1776
DB 21882 aggattatcatvggtcctttatagatgagtggtgtgtaattcaacglttcttaaccaa 21941
QY 1777 agagatcgggagagaacatagagaacttaaaacaggaagatgagaataactatgca 1836
DB 21942 gaacagcgtttatgattccgttaataatccaabaatttgaaagltccgtatctatctgt 22001
QY 1837 actgactagcctctagaagactgtagtctacgaagcttaacagatgataactaatcttga 1896
DB 22002 acggaatttgccctcagaaggtcttgatatacccaagattgaggtgttatataactagat 22061
QY 1897 ttccacggaalatitgaagaatacgtacacagaaatagggcgacaggaagagagag 1956
DB 22062 atgcacaagatlatagatctactactcatagagtggtgtcgtacgcgcagagctgtgtag 22121
QY 1957 actggtgttccattac 1973
DB 22122 gaagtgctgtccgtcac 22138

RESULT 13
US-09-318-443-7
/ Sequence 7, Application US/09318443
/ Patent No. 6197947
/ GENERAL INFORMATION:
/ APPLICANT: Helmslein, Daniel C.
/ APPLICANT: Hemmati-Brianlou, Ali
/ TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIII, AND METHODS OF USE
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: 600-1-211 N
/ CURRENT APPLICATION NUMBER: US/09/318,443
/ CURRENT FILING DATE: 1999-05-25
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: Patentn Ver. 2.0
/ SEQ ID NO 7
/ LENGTH: 1682
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-318-443-7

Query Match          4.3%; Score 101.6; DB 4; Length 1682;
Best Local Similarity 47.3%; Pred. No. 1.7e-16;
Matches 339; Conservative 0; Mismatches 374; Indels 3; Gaps 1;

QY 1257 tgtgtgtggaatagaatgaacaataagaagcttaaaaagtgtagatcataat 1316
DB 637 tgaaggcaccaatgttgcggagacatcagaagctgtgattcagacagcatgtgtgcgc 696
QY 1317 tgaactcccggaagaatgaaatgtaatctgcaaatgaaatlaactcgtcgaactcgaagaat 1376
DB 697 gggacatccagggcggtgtttttgatgatgtcgtcgcaagaagcctaagaacagtgat 756
QY 1377 aacctactgtgtttatagaatgaagacagaagaatgtgtgacatgaggaatttgaaccccgat 1436
DB 757 caaaatgttgttttgatgaaagtgtatgaaatgttgaatlaaagtttcaaaagacgat 816
QY 1437 aatgaagatttgtttagatgtgtgcgccagataagcagagactgaatgaacagatgtacatg 1496
DB 817 ttacgatagtatacaggtacocgtcctccacgacacaaagtggtgttccatcagtgacagct 876
QY 1497 gctcatatcagttcatgcgcctcgacacattcttcttgaagaagcaaatgattgtctgt 1556
```

```
Db 877 gccacagagatctcggaga tgaacacaagltca tgaacgccaatccgcact--t 933
Oy 1557 tggtaacttggaactagttgctgtaagttcagtgagcaaaatataatgttaaccgca 1616
Db 934 ggtgaaacggtatgaattgactctggaagcaccagaacatttctggtgcaatggaag 993
Oy 1617 ggaagagaatltgagltacatgtcaaacctttctacaagatgtatgtatccacagacaagt 1676
Db 994 ggaagatgtgaaatltgacactctgtgtgaccttaagacacactgacactgaactgagc 1053
Oy 1677 catgtctctgttctcgaaaagctgttcggatgaactacttaacagtgactaactattg 1736
Db 1054 ggtcatctcttcgcaacccaagaaggtgtgactgctgcagcagaagaatgtgaggaagc 1113
Oy 1737 aaatatacagtagagltctctgtcattgagatagagaagacagagatcgggagaaagcat 1796
Db 1114 caacttcaactgtatccctcaatgtcattgtgagacatgtcccaagaaggggagatccat 1173
Oy 1797 agagaactttaaacaagcagaatgtgagaaactaattgcaactgactatgaccttagagg 1856
Db 1174 gaaggaagttccggtcggggcgccagcgagtgctatttctacagatgtctggccaggg 1233
Oy 1857 actgtatgtccatgaacgttacaatgtctataatlttgaacttccacggaatatitgaag 1916
Db 1234 gttggaatgtccctcaggtgtccctcatcattacatagatcctccataatacagaagatt 1293
Oy 1917 atacgtacacccaataggggcgacaggaagcaggaagactgtgttccatt 1972
Db 1294 gtactaacacagaatttgggagatcaggtctgaacggcggaaggtgtgtgcatt 1349
```

```
RESULT 14
US-09-318-443-5
; Sequence 5, Application US/09318443
; Patent No. 6197947
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briyanlou, Ali
; APPLICANT: Weinstein, Daniel C.
; TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIIII, AND METHODS OF USE
; FILE REFERENCE: 600-1-211 N
; CURRENT APPLICATION NUMBER: US/09/318.443
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-318-443-5
```

```
Query Match 4.2%; Score 98.4; DB 4; Length 1536.
Best Local Similarity 47.1%; Pred. No. 1e-15;
Matches 337; Conservative 0; Mismatches 376; Indels 3; Gaps 1;

Oy 1257 tgggtgtggaataagatagaaacaaatagaaagacttaaaaaaggtatgatatcataat 1316
Db 496 tggaggaaccaatgttggcgaggacatcagaagcttgattacagacagcatgttgcgc 555
Oy 1317 tgaaccccggaagaattgaatgactgcaaatgagtaacttggtaactctgaagaatat 1376
Db 556 gggcaaccgagggcgtgttttggatatattctgcgaagaagccttaaggacaagtgcat 615
Oy 1377 aaactacttggtttatgatgaagcagaagaatgttgacaatggatattgaaccccat 1436
Db 616 caaaatgtgtgttttgatgaaagcgaatgaatgttgaataaagtttcaaaagacagat 675
Oy 1437 aatgaagatttgtatagatgtgcccagatagggcagacagattatcacagactgac 1496
Db 676 ttaagatgatataggttaccttgcttcagcacaacaggtgtgttcatcagaatgacagct 735
Oy 1497 gctcatcagttcatcgcttcgcacaactcttatttgaagaacaaatgatatgtctatgt 1556
```

```
Db 736 gccacagagatctcggaga tgaacacaagltca tgaacgccaatccgcact--t 792
Oy 1557 tggtaacttggaactagttgctgtaagttcagtggaagcaaaatataatgttaaccgca 1616
Db 793 ggtgaaacggtatgaattgactctggaagcaccagaacatttctggtgcaatggaag 852
Oy 1617 ggaagagaatltgagltacatgtcaaacctttctacaagatgtatgtatccacagacaagt 1676
Db 853 ggaagatgtgaaatltgacactctgtgtgaccttaagacacactgacactgaactgagc 912
Oy 1677 catgtctctgttctcgaaaagctgttcggatgtactatcaatgaaatgtgactaacttgg 1736
Db 913 ggtcatctcttcgcaacccaagaaggtgtgactgctgtcagcaggaatgtgaggaagc 972
Oy 1737 aaatatacagtagagltctctgtcattgagatagagaagacagagatcgggagaaagcat 1796
Db 973 caacttcaactgtatccctcaatgtcattgtgagacatgtcccaagaaggggagatccat 1032
Oy 1797 agagaactttaaacaagcagaatgtgagaaactaattgcaactgtatgaccttagagg 1856
Db 1033 gaaggaagttccggtcggggcgccagcgagtgctatttctacagatgtctggccaggg 1092
Oy 1857 actgtatgtccatgaacgttacaatgtctataatlttgaacttccacggaatatitgaag 1916
Db 1093 gttggaatgtccctcaggtgtccctcatcattacatagatcctccataatacagaagatt 1152
Oy 1917 atacgtacacccaataggggcgacaggaagcaggaagactgtgttccatt 1972
Db 1153 gtactaacacagaatttgggagatcaggtctgaacggcggaaggtgtgtgcatt 1208
```

```
RESULT 15
US-09-318-443-1
; Sequence 1, Application US/09318443
; Patent No. 6197947
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briyanlou, Ali
; APPLICANT: Weinstein, Daniel C.
; TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIIII, AND METHODS OF USE
; FILE REFERENCE: 600-1-211 N
; CURRENT APPLICATION NUMBER: US/09/318.443
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Xenopus laevis
; US-09-318-443-1
```

```
Query Match 4.1%; Score 96.4; DB 4; Length 1245;
Best Local Similarity 46.3%; Pred. No. 3.1e-15;
Matches 479; Conservative 0; Mismatches 531; Indels 24; Gaps 4;

Oy 961 gaagtatggaacacttaaaaagcaggtttcaaaagccagacacacatttcagtcacg 1020
Db 151 gaccttcgtagggcatctatgtcttatgtgatttggaaacacactggcctacacagaag 210
Oy 1021 gcatggccatgtgttgcaaggaatagatcttatagagtagcaggaactgtgaacagg 1080
Db 211 gcaatcaagcatcatcaataaagaaggaatgtatgtcgcacaatcacacactgtgtacagc 270
Oy 1081 aagacatgtgtatttaatgtcgtgagttatcatcgtgctctcaacccagccttaaa 1140
Db 271 aaacagcaacttttgtgttctgtctcaggtgttggatatcgaactcg----- 323
Oy 1141 gttcaaggaatagaccggcagatgttacttaactccactcgggaattagacattcaa 1200
Db 324 -----tgaacccaagccttgattttagcaccaccaccaaagattagacagcga 372
```

QY 1201 gt--agagagagaatcgttcgaataatca-pataaaggcttcggagtggttcgtat 1257
Db 373 atccagaagtgctcgtcgttcgtgggacacatgaatgctgacgtgcgtgtatc 432
QY 1258 ggtggcgaataaataagaatgaacaataaagaagcttaaaaaggctgaataatc 1317
Db 433 ggaagcacaataatcgtgaagaagatataccgaataatggaatataatgagacgttcgtc 492
QY 1318 gcaaccccggaagatctgaatcgtcaaatgaatgaacttcgtcaatcgtgaataata 1377
Db 493 ggaacacacgaagcgctgttcgtatcgttcgcacgcgaagatctaaagaatccggtccatc 552
QY 1378 acccaactggcttctagaatgaagaagaagaatgctgacacgtggatctgaaccacata 1437
Db 553 aaaaatgtaatgctgaatgaatgaatgaatgaatgctgaatgaatgctgaatgaatgaat 612
QY 1438 atgaagaatcttcgtgaatgctgacgaatgaatgaatgaatgaatgaatgaatgaat 1497
Db 613 taigtatgaatacgaatcgtcgtccacgaacacgaatgctgttaatacgaatgacacgtc 672
QY 1498 cctcaatcagctcaatcgcctcgcacaaatcttaattgaagaacaaatgaatgctatgct 1557
Db 673 ccacatgaataatccctggaataatgaacaaatgaatgaatgaatgaatgaatgaatga 732
QY 1558 ggtacatggaatcgaatcgtcgtgaatgaatgaatgaatgaatgaatgaatgaatga 1617
Db 733 aaacgtgaatgctgacacgtgaatgaatgaatgaatgaatgaatgaatgaatgaatga 792
QY 1618 gaagaagaatggaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1677
Db 793 gactgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 852
QY 1678 atgtcctcgttcctcgaagaatgctgacgaatgaatgaatgaatgaatgaatgaatga 1737
Db 853 atctcctcgaacacacaaagaagaatgct--agatcgtgactgacgaagaagaagaaga 909
QY 1738 aataatacgaatgaatgctcgtcgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1797
Db 910 aatcacaagatcctcgtcgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 969
QY 1798 gagaactttaaaacgaagaagaatgaatgaatgaatgaatgaatgaatgaatgaatga 1857
Db 970 aaagaatcctcgtcgtcgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1029
QY 1858 ctgaatcgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1917
Db 1030 ctggaatgctcgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1089
QY 1918 taatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1977
Db 1090 taatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1149
QY 1978 ctgaatgaatga 1991
Db 1150 gtcagaatgaatga 1163

Search completed: July 25, 2002, 12:19:07
Job time: 10947 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run On: July 25, 2002, 09:08:45 ; Search time 2465.67 Seconds

(Without alignments)
12945.893 Million cell updates/sec

Title: US-09-923-831-42

Perfect score: 2365

Sequence: 1 ttgtgacccagctcgtgaccc.....aaaaaaaaaaaaaaaaaa 2365

Scoring table: IDENTITY_MUC

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: em_estbta.*
2: em_esthum.*
3: em_estlin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hlc.*
9: 9b_estl.*
10: 9b_estl2.*
11: 9b_hlc.*
12: 9b_gss.*
13: em_gss_hum.*
14: em_gss_huv.*
15: em_gss_pln.*
16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	711	30.1	729	9	A1378017 t667f02.x
2	664	28.1	698	9	A1521173 th58d04.x
3	653.8	27.6	674	10	B6724056 602697033
4	638.6	27.0	742	10	B6772793 602720972
5	622.4	26.3	787	10	B6907039 60149866
6	618.8	26.2	639	10	B6724057 602697035
7	588.4	24.9	599	9	A1637624 t10c11.x
8	498.4	21.1	508	9	AL043701 dK2p434k
9	448.2	19.0	453	9	AA883800 al59d05.s
10	404	17.1	405	9	AA948168 cQ3b10.s
11	398.2	16.8	403	9	AA593728 hq21h02.x
12	397.8	16.8	564	10	AW341446 h009d07.x
13	387.2	16.4	404	9	B6193072 244514.MA
14	385	16.3	660	9	A1638049 t06f03.x
15	354.4	15.0	543	10	B6037235 B6037235
16	354.4	15.0	543	10	BE756114 210165.MA
17	330.6	14.0	679	9	AL637874 AL637874

18	330.4	14.0	504	10	B1066843
19	311	13.2	574	10	BE895251
20	296.8	12.5	568	10	BC089777
21	285.6	12.1	563	12	AO707919
22	247.6	10.5	562	10	BE367159
23	243.8	10.3	247	9	BE144935
24	238.6	10.1	564	10	B070471
25	231.4	9.8	809	10	B1464283
26	229.4	9.7	763	12	BH369602
27	218.8	9.3	543	10	BE575646
28	215	9.1	859	10	B1560190
29	210.4	8.9	231	10	BE377169
30	197.2	8.3	406	9	AA781181
31	197	8.3	861	12	CNS06D0
32	195.8	8.3	980	9	AL518383
33	190.8	8.1	996	12	CNS0722B
34	188.8	8.0	902	10	BM453752
35	188.8	8.0	1047	10	BM477615
36	188.8	8.0	1053	10	BM450043
37	188.8	8.0	1103	10	BM464068
38	184	7.8	799	10	BE607995
39	182.6	7.7	825	10	B1250982
40	181.4	7.7	664	9	AV705304
41	181	7.7	801	9	AV136032
42	180.4	7.6	827	12	A2667472
43	177.2	7.5	365	9	A1586361
44	176.2	7.5	735	9	AU126255
45	174.2	7.4	743	9	AV716479

ALIGNMENTS

RESULT 1
LOCUS A1378017/c
DEFINITION t667f02.x1 Soares_NFL_T-GBC_S1 Homo sapiens cDNA clone
IMAGE:2091771 3' similar to TR-01370 O13370 SUPPRESSOR OF
UNCONTROLLED MITOSIS. ; mRNA sequence.

ACCESSION A1378017
VERSION A1378017.1 GI:4187870
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 729)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

COMMENT This clone is available royalty-free through LINT; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1048 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 440.

FEATURES

1..729
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:2091771"
/clone_id="Soares_NFL_T-GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: p773D-Pac (pharmacia) with a modified polylinker. Site_1: Not I. Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized libraries (fetal lung, NbHL19W, testis NHT, and B-cell NCI-CGAP-CCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made

from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-727319, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo. " 2 others

BASE COUNT 196 a 170 c 105 g 256 t

ORIGIN

Query Match 30.1%; Score 711; DB 9; Length 729;
Best Local Similarity 98.4%; Pred. No. 4.4e-97;
Matches 717; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 1615 gaggaagaagaatgagtcacatgcaaaccttctcagagatgctatcaccagacaaa 1674
DB 729 GAAGAAGAGAAATGAGTACACTGCAAACTTTCTCAAAAGATGTCATCAAGACANA 670
OY 1675 gtcatgtctgtctctcgaagaaagctgttgcgaatcacttcaagtgactactt 1734
DB 669 GTCAATTGCTCTGCTTCTCGAANAAGCTGTGGGATCCTATCAAGTGACCTTAATCTT 610
OY 1735 ggaataatcatagtagtctctcgaatgagatgaaacagagagatccggagaaaga 1794
DB 609 GGAATATATACATAGTACAGTCTCTGATGAGATAGAGAACAGAGATCGGAGAAAGCA 550
OY 1795 tttagagactttaaaacagagcaagtgagaacttaattgcaactgactagcctcaga 1854
DB 549 TTAGAGNCCTTTAAACAGGCAAGTGAATTAATTAATTAATTAATTAATTAATTAATTA 490
OY 1855 ggaactgagtcagtagaagttacacatgcttaatttgaacttccacaggaataatgaa 1914
DB 489 GGACTTGATGTCATGACCTTACACATGCTCTATTAATTTGACTTTCACGGAATATGAA 430
OY 1915 gaatacgtacacagaaatgagcgacgagagagagagagagagagagagagagagag 1974
DB 429 GAATACCTACACGAAATAGCGGACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 370
OY 1975 acttgcagtagaaatgagtgagagtgctcctcgaatgattataatctcggaaagaga 2034
DB 369 ACTTGACTGAAATGATTTGAGAGGTGCTCTGAAATTAATTAATTAATTAATTAATTA 310
OY 2035 aatcagaatattccagagagagcttgatcaatgctgagagaggtttgagagacatacag 2094
DB 309 AATCAGAGTATTCAGAGAGAGCTGTATCAATGTGCTGAGGTTTGAGGACATCAACG 250
OY 2095 aaaaaggaatggaagaaagaaatggaagagcctcagagagagagagagagagagag 2154
DB 249 AAAAGGGAATGGAAGAAATGGAAGAACCTCAAGAGAGGCCCAAGAGATTCATTAA 190
OY 2155 tgcctcgtactagtgaggttagagaattcaagaatttttgaagaatataagtaagaagaa 2214
DB 189 TGCTCTGCTACTAGTGGGGTGAAGATTTCAAGATTTTGAATATATAGTAAAGAGAA 130
OY 2215 gtattgacactgttgcagatgataagagacagcctgatttgactgattcttaataaat 2274
DB 129 GTATTGGACATGTTGGCACTATGAAGACCGGACCTGATTTGACTGATTTTAAATAT 70
OY 2275 agtctcgaagaatataagatccagtgcttatacttcttaataaataagaagattt 2334
DB 69 AGTGTTGAAATATATGAAATCCAGTGTATTACTTCTTATAAATAATGAGATATT 10
OY 2335 aaacttaa 2343
DB 9 AAACCTTAAA 1

RESULT 2
A1521173/c
LOCUS A1521173 698 bp mRNA linear EST 13-APR-1999
DEFINITION t58404.xl NCI CGAP Kid1 Homologous cDNA clone IMAGE:2122471 3'
similar to TR:013370 O13370 SUPPRESSOR OF UNCONTROLLED MITOSIS. ;
mRNA sequence.
ACCESSION A1521173
VERSION A1521173.1 GI:4435308

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 698)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnln.gov/brp/image/image.html
Insert length: 2152 Std Error: 0.00
Seq primer: -400P from G1bco
High quality sequence stop: 424.
Location/Qualifiers

FEATURES
source 1..698
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2122471"
/clone_1lb="NCI-CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pRT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
plasmid DNA from the normalized library NCI-CGAP_Kid1 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1322376-1323911, 1456007-1456775, and
1500352-1502835). Subtraction by Bento Soares and M.
Fatima Bonaldo. "

BASE COUNT 182 a 158 c 93 g 265 t

ORIGIN

Query Match 28.1%; Score 664; DB 9; Length 698;
Best Local Similarity 97.8%; Pred. No. 4.7e-90;
Matches 673; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1678 atgtcttcgttctcgaagagctgttgcgagatcactatcaagtgacctaacttga 1737
DB 698 ATGTCTTCGTTTCAAAAAGCTGTGCGGATCACTAATCAAGTACTTAATACGAA 639
OY 1738 aatataatgaatgagtcctcgaatgagaataagaacagagagagagagagagagat 1797
DB 638 AATATATCAGTAGAGTCTCTGATGAGATGAGACACGAGAGATCGGGAAGCATTA 579
OY 1798 gagaactttaaaagagcaagtgagaataactaattgacactgactagactcactaga 1857
DB 578 GAGACTTTAAAGAGCGCAAGAGTGAATATCAATTCAGATGATCAAGCTCTAGAGA 519
OY 1858 cttagatgcatgacgtttacacatgcttaattttagacttccagagagatattgaaga 1917
DB 518 CTTAGATGTCAGTAGAGTCTCTGATGAGATGAGACACGAGAGATCGGGAAGCATTA 459
OY 1918 taagtacacagaaatgagcgacgagagagagagagagagagagagagagagag 1977
DB 458 TAGCTACACCGAATAGGCGCACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 399
OY 1978 ttgacagagaatgagtgagtggtgtgctctgagatgattatctcggaaagagaat 2037
DB 398 TTGACTAGGAATGATGAGGTGCTCTGATGATTAATTAATTCGGAAGAGCAAT 339
OY 2038 cagagatattccagagagagcttgatcaatgagtgagaggtttgagagacatacaga 2097

```
|||||
Db 338 CAGAGTATTCAGAGAGCTGTATCAATGGCTGAGAGGTTAAAGCACATCAGACAAA 279
QY 2098 agggaaatgaaagaaatgaaagacctcaagaaagcccaaaagttcataatgt 2157
Db 278 AGGGAATGMAAAGAAATGMAAGACCTCAGAGAGGCCCAAGAACTTTCAATTAATGT 219
QY 2158 ctctgactagtggggtagaagattctttagaaatagtagaagacagaagta 2217
Db 218 CTTCTGTACTAGTGGGTAGAGAAATCAAGATTTTAAATATAGTAACGCGAAGTA 159
QY 2218 ttgagacatgtgacagtagaagagaccgagctgattacagattcttaaaatag 2277
Db 158 TTGGACATGTTGGCAGTGTGAAGAGACCGACATGTTTACATGATTTTAAATATAGT 99
QY 2278 gttgaaatatagaaaccagtgcttatacttctttaaataaataagaatataa 2337
Db 98 GTTTGAAATATAGAAATCCACTGTTTATCTTTCTTTAAATATAGATTTAA 39
QY 2338 cttaaaaaaaaaaaaaaaaaaaaaa 2365
Db 38 CTTGGAATAAAGAAAGAAAGAAAGAAAGAA 11

RESULT 3
Bg724056 674 bp mRNA linear EST 08-MAY-2001
LOCUS 60269703F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4829072 5',
DEFINITION mRNA sequence.
ACCESSION Bg724056
VERSION Bg724056.1 GI:14003243
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 674)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0748 row: a column: 09
High quality sequence stop: 674.
Location/Qualifiers
1..674
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4829072"
/cclone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.2 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 219 a 137 c 182 g 136 t
ORIGIN

Query Match 27.6% Score 653.8; DB 10; Length 674;

Best Local Similarity 99.6%; Pred. No. 1,6e-88;
Matches 666; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 183 gtcgagcgccctctcttggaacaa tgcacacagagagctcccaagccctac 242
Db 6 GCGGAGCCGCCCTCTTGTGAAATGTCACACAGAGAGACTCCAGGCCCTAC 65
QY 243 gtcggtgctgctagtcgagagctcgagctgccccgagccgagagagcgccg 302
Db 66 GTGGGCTCTTCTAGTGGCGAAGCTGCAATGTCCGAGCCGCAAGAGAGGCGGCG 125
QY 303 ggaagagctgaatcgaacaggtccgtagaagataatgctcgagagagtgctgtag 362
Db 126 GGAGAGCTTGAATGCAACAGCTCTTACGAGATATAGTGTGCGAGAGGTGTGCGGAG 185
QY 363 aggcacctctagcgcccgagagcggtgagcgctggtcaagaggaactgcgcgtgt 422
Db 186 AGCGACCTCTAGGCGCCCGGAGGACGTGGCGGTGACAGAGAACTCCGCTGTGTT 245
QY 423 tgcctgaaagacacttgcttgagcggtatcgtgctgctgctgctgctgctgct 482
Db 246 TCGTTTGAAGAGCCACTTGTGTGGCGGTATGCGTGTGCGGTCAAAATTAAGAA 305
QY 483 tatcaaaagtacaacaacaacccaatccaataataaagaacaacgaatcattgt 542
Db 306 TATCAAAAGTACAAACAAACACACATCCAAATATACAAAGCAACGAAATCAATAGT 365
QY 543 caaaatcttgagcaagcaatgcgaacgaagcaaaagcagtagatagacaatttgt 602
Db 366 CAAATATTTTGGCAGCAGGCAATGCAAAAGCAAAAGCAGTAGTAGCAATTTTGT 425
QY 603 taaaagctagaagaataattacaattcagaatgaggaattgatactgcatccaactc 662
Db 426 TAAAGCGTGAAGAAATTTACAAATTCAGAAATGCGGAATGATGCAATCAACCTTC 485
QY 663 gtttgagaag-agaatgaagacagataacaaatgttgttgagagagatcgccattgag 721
Db 486 TGTGGAAAGAGATGAGACACAGATTAATGTTGTCAGAGATGCGCATTTAG 545
QY 722 attggaatcaaatagagagaggttgaaatgagcaaaaacaaagtgagagattac 781
Db 546 ATTGGATTCAAATTAGAGAGAGAGTTTGAATATGCAAAAGATGGAGATTTAC 605
QY 782 caaccaataagaaaaacttctaagaagctcactgcacaagtgccatgtaaaagtag 841
Db 606 CACCAATTAAGAAATCTTTATTAAGAGTCCACTGCCACAAAGTGCATGTAAGAATAG 665
QY 842 aagcagata 850
Db 666 AAGCAGATA 674

RESULT 4
Bg72793 742 bp mRNA linear EST 15-MAY-2001
LOCUS 602720972F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4837888 5',
DEFINITION mRNA sequence.
ACCESSION Bg72793
VERSION Bg72793.1 GI:14083446
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 742)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
```



```
QY 1854 aggaacttgatgtccatgacgttacaatgtctataattctgacttccacgaatattga 1913
|||||
Db 301 AGGACTGTATGTCATGACGTTGACATGTCATATATTTCATCTTCCACGGAAATTTGA 360
QY 1914 aqaatacgttacacccaatagggcgacgggaagacgaagagacgttggttccattac 1973
|||||
Db 361 AGAATACGTACACCGAATAGGGCGACGGGAAGACAGGGAGGACTGTGTTCATTCAC 420
QY 1974 aacttgactagaatattgagaggctgcccgtgacttgattataattctggaaagagc 2033
|||||
Db 421 AACTTGACTAGAAATGATTGGAGGGTGCCTGATTAATTGATTATTTCTGGAAAGAC 480
QY 2034 aaatacagatattccaaagagacgttgatcatatgctcgaagaggtttgagacatacaag 2093
|||||
Db 481 AATATCAGATATTTCCAGAGAGCTTTATCAATGGCTGAGAGGTTTGAGGCACATCAACG 540
QY 2094 gaaaagg---aaatgaaagaaatgaaagacctcaaggaagcccaagaagtctc 2149
|||||
Db 541 GAACACAGGACATGGAACCGAATAATGAAAGACCTCAAGAGAGGCCCAAGAAC -TTC 599
QY 2150 attaatgctctctgactagtgagagag---aattcaagaatttttagaataata-gt 2205
|||||
Db 600 ATTTAATGCTCTGTACTAGTGGGGCTAGAGAAATTCACGATTTTAAATAATTAAGGT 659
QY 2206 aagacag-aagtatcgacatgttcgcaatgaa--gagacggactgacttgactgat 2262
|||||
Db 660 AAGACAGAAATATCGACATGTTGGCTGACTGACGAGGCCGAGCCGTTGACCTGAT 719
QY 2263 tcttaaaatagtagtcttgaataatagaaatcgaagtggttttaacttcttcaataaa 2322
|||||
Db 720 CCTTACACTATATGTTTCG-AACCTTGAATCAAGGGTGTATATCTTCTTCATCAAAA 778
QY 2323 atag 2326
||
Db 779 TGAG 782

RESULT 6
Bg724057 639 bp mRNA linear EST 08-MAY-2001
LOCUS 602697035F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4829168 5',
DEFINITION mRNA sequence.
ACCESSION Bg724057
VERSION Bg724057.1 GI:14003244
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 639)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strussberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10748 row: e column: 09
High quality sequence stop: 636.
Location/Qualifiers
1..639
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4829168"
/clone_11b="NIH_MGC_97"
/lab_host="DH10B"

FEATURES
Source
```

```
Query Match 26.2%; Score 618.8; DB 10; Length 639;
Best Local Similarity 99.5%; Pred. No. 2,7e-83;
Matches 631; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 183 gtctgagcgccctcttctggaacaatgtccacacgaagagaggtcccaagcctcac 242
|||||
Db 6 GGGGAGCGCGCCCTTCTTGGAAACATGTCCACACAGGAGAGCTCCCAAGGCTCTAC 65
QY 243 gtgggtcgtctgactcgagagctcgacagtgctccgagcgagagagagcgagc 302
|||||
Db 66 GTGGGTGCTTGTAGTCGCGCAAGCTGACAGTGTCCCGAGCGCCAGAGAGAGCGCGC 125
QY 303 gggagagttgaatcgaaacaggtcctcgagagatagtgctcgacagagagtgctcgag 362
|||||
Db 126 GGAAGAGTTAATCGAACAGGTCTGAGAGGATTAAGTGTGCGCAGAGGTGTCTCTGGAG 185
QY 363 aggaaccttaggcgcccgagagcgctgagcgctgagtcacgaaggaactgacgctgtct 422
|||||
Db 186 AGGCACCTTAAGGCCGCCGAGAGAGGTGGCGGTGTCAGAGAGAACTGCGCTGTGTT 245
QY 423 tgccttgaagagcgacctgtgtgagcggttaactcgtgtgtgtgttcaaaaataagaa 482
|||||
Db 246 TGCTTTGACAGCCACTTTGTGTGGCGGTAACTGCTGTGTGTGCTGCTGCTGCTGCTG 305
QY 483 tatcaaatgtaacaacacacacacacacacacacacacacacacacacacacacacac 542
|||||
Db 306 TATCAAAAGTACACAAACACCAACATCCAAATTAATACAAACAAACCAATCATTAAGT 365
QY 543 caaatcttggagcaagcgcaatgcaagcaagcaagcaagcaagcaagcaagcaagca 602
|||||
Db 366 CAATAATTTTGGCAGCAAGCCATATGCCAAGCAAAAGCAAAAGCAATTTTGT 425
QY 603 taaaagctagaagaanaattacaatlcagaatcggaattgatactgcattccaacctc 662
|||||
Db 426 TAAAAAGCTAAGAAATAATTAACAATTCAGAAATCGGAATGTACTGCAATCCAACTTC 485
QY 663 tgttggaagaagatggaagcacagatacaatgctgtgtgagagagatgagcaatgata 722
|||||
Db 486 TGTGGGAAGAATGGAACACACAGATTAACAATGTTGTTCAGAGAGATTCGGCATTTGATGA 545
QY 723 ttggagatcaaatgagagagaggtctgaaatggcaaaaacaaatgggacagattacc 782
|||||
Db 546 TTGGGATCAAAATTAAGAGGAAAGTTGAATAG-AAAAACAAAGTGGGAGATTTAAC 604
QY 783 accaatgaagaaaactttataagaggtccact 816
|||||
Db 605 ACCAATTAAGAAAACCTTTTATAAGAGTCCACT 638

RESULT 7
A1637624 599 bp mRNA linear EST 14-DEC-1999
LOCUS A1637624
DEFINITION t10c11.x1 NCI_CG66 Homo sapiens cDNA clone IMAGE:2240372 3',
similar to TR:059857 059857 MOC2 RNA HELICASE. [1] ;, mRNA
sequence.
ACCESSION A1637624
VERSION A1637624.1 GI:4689858
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 599)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILMIL at: www.bio.linn.gov/bbrp/image/image.html
Insert Length: 495 Std Error: 0.00
Seq primer: -400p from Glbco
High quality sequence stop: 444.
Location/Qualifiers

FEATURES
source

1..599
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2240372"
/clone_lib="NCI-CGAP_GC6"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; plasmid DNA from the normalized library NCI-CGAP_GC4 was prepared, and as circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonides 1257096-125851, 1469064-1470983, and 1473592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 159 a 141 c 78 g 221 t
ORIGIN

Query Match 24.9%; Score 589.4; DB 9; Length 599;
Best Local Similarity 99.0%; Pred. No. 6, 9e-79;
Matches 593; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1753 tctctgcaaggagatagagacagagatcgaggagaagcattagagaacttaaacaa 1812
|||||
DB 599 TCTCTGCATGAGATAGAGAACAGAGATCGGCAAAACATTAGAACTTTAAACA 540
OY 1813 ggcaaaagtcgaaatactatctgcaactgacttagcctctagaggacttgatgccatgac 1872
|||||
DB 539 GGCAGAGTGAGAACTACTATTTCAGACTGATTCAGCCCTGAGAGACTTGATTCAGTAC 480
OY 1873 gtacacatgctataaatttgacttccacggaatatgaaagaatgctacacgaata 1932
|||||
DB 479 GTTACACATCTCTAATAATTGACTTCCACGGAATTTGAAATAATACGTACACGAATA 420
OY 1933 ggagcgacaggaagacagagagagctggtgttccattcaacttgacttgactgaatgat 1992
|||||
DB 419 GGGCGCACGGGAAGACGAGAGACTGTGTTCCTTCACTTACACTTTGACTGAAATGAT 360
OY 1993 tggagggatgctctgcaattgataatcttgaaaagacaaatcagaatctccagag 2052
|||||
DB 359 TGGAGGGTGCCTCTGAATTTGATTAATTTCTGGAAAGACCAATAGAGATTTCAGAG 300
OY 2053 gagcttgatcaatgctcgaagggttgaaggacacatacaggaagaagggaagtgaaga 2112
|||||
DB 299 CAGCTTGATCAATGCTGAGAGGTTTAAAGCACATACACAGAAAAGGGAATGGAAGA 240
OY 2113 aaaaatgaaagaccctcaagaaagcccaagaagttcattcaatgctcttgcataatg 2172
|||||
DB 239 AAAATGGAAGACCTCAAGAGAGGCCCAAGAAATTTCATTAAATGCTCTGTAATGAG 180

OY 2173 gctagagaaattcaagaattttttagaataatagtaagacaagaatgattggacatgtcgca 2232
|||||
DB 179 GGTAAAGAAATTCAGAAATTTTGAATAATTTACTTAAGACGGAAGATTTGGACATGTTGGCA 120
OY 2233 gattgaagaagaccggaactgattgactgacttcttaataataatagtgcttgaaataatga 2292
|||||
DB 119 GATGAAAGACCGGACTGATTTGAATGATCTTAAATAATATGTTGAAATAATATGA 60
OY 2293 atccagtgcttacttcttcttaataaataagaagatcattcaacttaaaaaa 2351
|||||
DB 59 ATCCAGCTGTTTATTAATCTTTCTTAATAAATAATAGAGATTTAAATGAAAAA 1

RESULT 8
AL043701/c 508 bp mRNA linear EST 29-FEB-2000
LOCUS
DEFINITION DKF2p434K0927.s1.434 (synonym: htes3) Homo sapiens cDNA clone
AL043701 DKF2p434K0927.3', mRNA sequence.
ACCESSION AL043701.1 GI:5423087
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Wiemann,S., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
EST (Bloecker, et al.)
Unpublished (1999)
Contact: Bloecker H

TITLE
JOURNAL
COMMENT

MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No RI sequence available.
This clone (DKF2p434K0927) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers

FEATURES
source

1..508
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2p434K0927"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 143 a 117 c 71 g 177 t

Query Match 21.1%; Score 498.4; DB 9; Length 508;
Best Local Similarity 98.8%; Pred. No. 2, 9e-65;
Matches 502; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1824 aatacctaattgcaactgactcgaagctctgagagacttgatgcatgaacttacaatgt 1883
|||||
DB 508 AATACTAATTTGCAACTGATCTAGCCTCTAGAGACTTGATGCTGACGTTCACATGT 449
OY 1884 ctataatttgacttccacaggaatctgaagatactacaccgaaatagggcgacaggg 1943
|||||
DB 448 CTATTAATTTTGACTTCCACGGAATTAATTAATAATACGTAACGGAATAGGGCGCTCGGC 389
OY 1944 aagaagcaggagagactggtgttccattcaacttgacttgactgaatgattgaggggcttc 2003
|||||
DB 388 AAGAGCAGGAGACTGCTGTTTCCATTACAACTTGAATGACTGAATAATGATGAGGCTTGC 329

```
Qy 2004 cctcgaatcgaataatctcggaaagacaaatcagaatcattccagaagcctgtatc 2063
|||||
Db 328 cttctgaattgattatatttctgaaagacaaatcagagctttctcaagagcgtttatc 269
|||||
Qy 2064 aatgacctgaagagcttctgagcacatcaacggaaaaggaaatggaagaaatggaag 2123
|||||
Db 268 aatgacctgaagagcttttaagccacatcagcgaatggaagaaatggaagaaatggaag 209
|||||
Qy 2124 accccaaggaagcccaagaaagcttcattcaatctctctctgagcagtgaggtagagaatc 2183
|||||
Db 208 acctcagagagcccaagaaagctttcattcaatctctctgagcagtgaggtagagaatc 149
|||||
Qy 2184 caagatcttctgaataatagtaagacagaagatcttgacatgcttgagcagtagaaga 2243
|||||
Db 148 caagatcttctgaataatagtaagacagaagatcttgacatgcttgagcagtagaaga 89
|||||
Qy 2244 ccgagactgattcgaatcattccttaataatagtgcttctgaataatagtagatccagtgctt 2303
|||||
Db 88 ccgagactgattcgaatcattccttaataatagtgcttctgaataatagtagatccagtgctt 29
|||||
Qy 2304 tatctctcttcttaataataatagaagta 2331
|||||
Db 28 tatctctcttcttaataataatagaagta 1
|||||

RESULT 9
AAB83800 453 bp mRNA linear EST 04-JAN-1999
LOCUS A159d05.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
DEFINITION IMAGE:1461609 3' similar to SW:DB10_NICSY P46942 RNA HELICASE-LIKE
PROTEIN DB10. ; mRNA sequence.
ACCESSION AAB83800
VERSION AAB83800.1 GI:2993330
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 453)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone; similarity on wrong strand
Insert Length: 855 Std Error: 0.00
Seq primer: -40ml3 fwd. RT from Amersham
High quality sequence stop: 280.
Location/Qualifiers
1..453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1461609"
/clone_11b="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site:1: Not I; Site:2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBH19w, testis NH7, and B-cell
NCI-CGAP-GCB1) were mixed, and ss circles were used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 125 a 110 c 58 g 160 t

Query Match 19.0%; Score 448.2; DB 9; Length 453;
Best Local Similarity 99.3%; Pred. No. 9,6e-58;
Matches 450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1891 ttgacttccagaataatctgaagaatactgaacccgaatagggccacgggaagagca 1950
|||||
Db 453 ttgacttccagaataatctgaagaatactgaacccgaatagggccacgggaagagca 394
|||||
Qy 1951 gggaggaactgggttccatctacaaacttgactagaatgagtgagggttgctctgaa 2010
|||||
Db 393 gggaggaactgggttccatctacaaacttgactagaatgagtgagggttgctctgaa 334
|||||
Qy 2011 ttgactaatactcggaaagacaaatcagaagatcctcagagagccttgatcattgct 2070
|||||
Db 333 ttgattatatttctgaaagacaaatcagaagatcctcagagagccttgatcattgct 274
|||||
Qy 2071 gagaagcttgaagcacatcaacggaaaaggaaatggaagaaatggaagaccccaa 2130
|||||
Db 273 gagaagcttgaagcacatcaacggaaaaggaaatggaagaaatggaagaccccaa 214
|||||
Qy 2131 ggaagcccaagaagcttcattcaatctctctctgagcagtgaggtagagaatcagaatc 2190
|||||
Db 213 ggaagcccaagaagcttcattcaatctctctctgagcagtgaggtagagaatcagaatc 154
|||||
Qy 2191 tttaagaataatagtaagacagaagatcttgacatgcttgagcagtagaagaccggact 2250
|||||
Db 153 tttaagaataatagtaagacagaagatcttgacatgcttgagcagtagaagaccggact 94
|||||
Qy 2251 gatttgactgattccttaataataatagtgcttctgaataatagtagaactgcttattcctt 2310
|||||
Db 93 gatttgactgattccttaataataatagtgcttctgaataatagtagaactgcttattcctt 34
|||||
Qy 2311 tctttaataataatagtagatcttaactcaa 2343
|||||
Db 33 tctttaataataatagtagatcttaactcaa 1
|||||

RESULT 10
AA948168 405 bp mRNA linear EST 23-JUL-1998
LOCUS AA948168/c
DEFINITION OQ33P10.S1 NCI-CGAP_GC4 Homo sapiens cDNA clone IMAGE:1588123 3',
MRNA sequence.
ACCESSION AA948168
VERSION AA948168.1 GI:3109421
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 405)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
cDNA library preparation: M. Bento Soares, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 484 Std Error: 0.00
Seq primer: -40ml3 fwd. RT from Amersham.
Location/Qualifiers
1..405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1588123"
```

/clone.lib="NCI CGAP GC4"
/tissue.type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

BASE COUNT 114 a 100 c 48 g 143 t
ORIGIN

Query Match 17.1%; Score 404; DB 9; Length 405;
Best Local Similarity 100.0%; Pred. No. 4e-51;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1937 gcaaggaagagcagagagagactggtgttccatcaacttgaactagaatgttga 1996
DB 405 GCACGGAGAGACGAGGAGCTGTGTTCCATTACACTTGTGACTAGAAATGATTGGA 346
QY 1997 ggttgccctcgaattgaataatctcgaagaagcacaatccagaatccagaagagc 2056
DB 345 GGGTGGCCTCGAATTGATTATATTCTGAAACAGCAATCAGACTATTCCAGAGAGC 286
QY 2057 ttgttcaatgctcgtgagaggttggagcactcaacgaaaggaatggaagaaataa 2116
DB 285 TTGTATCAATGCTGAGAGGTTGAGCAATCAACGAAAGAAATGAAACAAAAA 226
QY 2117 tgaagaagacctcaagaagcccaagaagttcattatgctcctgaactagtgggta 2176
DB 225 TGAAGAAGCCTCAAGAGAGGCCCAAGATTCAATTAATGTTCTGTACTAGTGGGTA 166
QY 2177 gagaattcaagattttttagaataatagtaagacagaagatgagcagtggcagat 2236
DB 165 GAGAAATCAAGATTTTATGAAATATAGTAAGACAGAGATTGGACATGTGGCAGAT 106
QY 2237 gaagaagccgagactgattgactgattcttaataatagtgcttgaataatagaatcc 2296
DB 105 GAAGAGACCGGACTGATTGACTGATCTTAATAATATATAGTTGAAATATAGATCC 46
QY 2297 agtgtttatacttcttataataaataagaagatattaact 2340
DB 45 AGTGTTTATACCTTCTTTAATAATAATAGATATTAACTT 2

RESULT 11 405 bp mRNA linear EST 22-MAR-2000
AM593728/c hg21h02.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2946291 3'
LOCUS similar to TR:O15919 O15919 MITOCHONDRIAL DEAD BOX PROTEIN.; mRNA
DEFINITION sequence.

ACCESSION AM593728
VERSION AM593728.1 GI:7280986
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI CGAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
Image.llnl.gov/image/llnl/resources.shtml
Seq primer: -400p from Gibco.

FEATURES

source
Location/Qualifiers
1. 405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2946291"
/clone.lib="NCI CGAP GC6"
/tissue.type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; Plasmid DNA
from the normalized library NCI CGAP GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonoids
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 114 a 100 c 48 g 143 t
ORIGIN

Query Match 17.1%; Score 404; DB 9; Length 405;
Best Local Similarity 100.0%; Pred. No. 4e-51;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1937 gcaaggaagagcagagagagactggtgttccatcaacttgaactagaatgttga 1996
DB 405 GCACGGAGAGACGAGGAGCTGTGTTCCATTACACTTGTGACTAGAAATGATTGGA 346
QY 1997 ggttgccctcgaattgaataatctcgaagaagcacaatccagaatccagaagagc 2056
DB 345 GGGTGGCCTCGAATTGATTATATTCTGAAACAGCAATCAGACTATTCCAGAGAGC 286
QY 2057 ttgttcaatgctcgtgagaggttggagcactcaacgaaaggaatggaagaaataa 2116
DB 285 TTGTATCAATGCTGAGAGGTTGAGCAATCAACGAAAGAAATGAAACAAAAA 226
QY 2117 tgaagaagacctcaagaagcccaagaagttcattatgctcctgaactagtgggta 2176
DB 225 TGAAGAAGCCTCAAGAGAGGCCCAAGATTCAATTAATGTTCTGTACTAGTGGGTA 166
QY 2177 gagaattcaagattttttagaataatagtaagacagaagatgagcagtggcagat 2236
DB 165 GAGAAATCAAGATTTTATGAAATATAGTAAGACAGAACTATGTGGCAGAT 106
QY 2237 gaagaagccgagactgattgactgattcttaataatagtgcttgaataatagaatcc 2296
DB 105 GAAGAGACCGGACTGATTGACTGATCTTAATAATATATAGTTGAAATATAGATCC 46
QY 2297 agtgtttatacttcttataataaataagaagatattaact 2340
DB 45 AGTGTTTATACCTTCTTTAATAATAATAGATATTAACTT 2

RESULT 12 403 bp mRNA linear EST 31-JAN-2000
AM341446/c hd09d07.x1 Soares_NFL.T.GBC.S1 Homo sapiens cDNA clone
LOCUS IMAGE:2909005 3' similar to SW:DBJ3_YEAST P20447 PROBABLE
DEFINITION ATP-DEPENDENT RNA HELICASE DBP3.; mRNA sequence.

ACCESSION AM341446
VERSION AM341446.1 GI:6838072
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
I (bases 1 to 403)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
 Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps@email.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -400P from Gibco.

FEATURES
 source
 1..403
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2909005"
 /clone_1lb="Soares_NFL_T_GRC_S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NDHL19W, testis NHT, and B-cell
 NCL-CGAP GC81) were mixed and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."

BASE COUNT 114 a 97 c 46 g 146 t
ORIGIN

Query Match 16.8%; Score 398.2; DB 9; Length 403;
Best Local Similarity 99.3%; Pred. No. 2,9e-50;
Matches 400; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1940 cggagagagcagagagcgtggttccattacacatttgcacgaagaatcagatgagag 1999
 |||||||
 Db 403 CGGGAGAGACAGAGAGAGCTGGTCTTCCCTTACACTTTCAGAAATGATTTGGAGG 344
 |||||||
 Oy 2000 ttgacctgaattgattatatttcgaaagacaaatcagaatctccagagagcttg 2059
 |||||||
 Db 343 TTGCCCTTGAAATTGATTAATTTCTGGAAAGAGCAATTCAGATTCAGAGAGCTTG 284
 |||||||
 Oy 2060 tatcaatgactgaagaggtttgagggcacatcaacgagaaaggaatgaaagaaaatg 2119
 |||||||
 Db 283 TATCAATGGCTGAGAGGTTTTCAGGCACATCAACGAAAGGAAATGGAAGAAATGG 224
 |||||||
 Oy 2120 aaagacctcaagaaagcccaagaagttcatatcattcctctgtaagtgggtagag 2179
 |||||||
 Db 223 AAAGACCTTCAGAAAGGCCCAAGAGTTTCATTAAATGCTTCTGACTAGTGGGTAAG 164
 |||||||
 Oy 2180 aatcaagaattttagaataatagtaagacagaatcagatcagacatgttcagatgaa 2239
 |||||||
 Db 163 AATTCAAGATTTTAAAGAAATATAGTAAGACAGAAATATGACATGTGSCAGATAGAA 104
 |||||||
 Oy 2240 gagaccgagctgattgactgattccttaataatagtgcttgaatatagaatccagt 2299
 |||||||
 Db 103 GAGACCGAGCTGATTTGACTGATTTTAAATATATGCTTTGAAATATATGATTCAGT 44
 |||||||
 Oy 2300 gtttactactcttaataaataagaatatttaactaa 2342
 |||||||
 Db 43 GTTTTACTTCTTATTAATAAATAGAACTATTTAAATTA 1

RESULT 13
 BF193072 564 bp mRNA linear EST 02-NOV-2000
 LOCUS BF193072
 DEFINITION 244514 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BF193072
 VERSION BF193072.1 GI:11076441
 KEYWORDS EST.
 SOURCE pig.

ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE
AUTHORS Fahrenkrug, S.C., Fraking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, F.,
 Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, M.W.,
 and Keele, J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith RPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390

Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt-trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -mnscore 18
 and -mismatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCATGACGACG
 Plate: 71 Row: L Column: 19
 Seq primer: ATTTAGGTGACATATAG.

FEATURES
 source

1..564
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_1lb="MARC 2P1G"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from testis, ovary,
 endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 180 a 103 c 129 g 152 t
ORIGIN

Query Match 16.8%; Score 397.8; DB 10; Length 564;
Best Local Similarity 82.4%; Pred. No. 3e-50;
Matches 472; Conservative 0; Mismatches 92; Indels 9; Gaps 1;

Oy 901 gagaagacactatcccaatccatccagcatttgatgacgcttcaatgattactc 960
 |||||||
 Db 1 GAGAAAGCTCTTATTCCTATTCCTGCTGTAATTTGAGATGATTCAGTTCACCT 60
 |||||||
 Oy 961 gaggttatgagaataacataaagcaggtttcaaaagccaaacctatctcaagcaag 1020
 |||||||
 Db 61 GAAGTAGATGAATAACTTAATAAAGCGAGGTTTCAAAAAGCCAACTCAATTCAGTCACAG 120
 |||||||
 Oy 1021 gcatgcccattgctgctcaaggaatagatcttaagagtaagcagacatgaaacagga 1080
 |||||||
 Db 121 GCATGCCCAATCATTTACAGAGAAATAGATCTTATAGAGATGAGCCAGACTGAAACAGG 180
 |||||||
 Oy 1081 aagaatctgttatcttaagctcgtgattatctatctcgttcctcaaccgaccttaa 1140
 |||||||
 Db 181 AAGACATGTTCTTAACCTTAATGCTTGATTTATTAATGTTGACTACACACC-----T 231
 |||||||
 Oy 1141 ggtcaagaagatagaccggcagatgtcagttcctcaactcccatcgagggaatcagacctcaa 1200
 |||||||
 Db 232 GTAGCCAGAGATGGCCCGGCGCATGTAGCTTACACCACTGAGAAATGATGCTGCAG 291
 |||||||
 Oy 1201 gtaagaagagagatgtgcaaatatcatataaaggcctcggaggtgttgatataag 1260
 |||||||
 Db 292 GTGGAAGCTGAGTGTCTTAAGTATTCATTAAGAGCTTTAAAGGTGTGTGATATGCT 351
 |||||||
 Oy 1261 ggtggaatagagatgacaatagagaagcttcaaaaagggtgtagataatcagatcag 1320
 |||||||
 Db 352 GGTGAGATCGAGATGAGCAAAATAAAGACTTATTAAGAGGTGATATCATCATTCGA 411
 |||||||
 Oy 1321 actccggaagatgaatgataatcgcgaatggaacttgtaactcgaagaatataacc 1380
 |||||||
 Db 412 ACTCTGGAGGCTCAATGATCTGCAAAATGAATTAATCTTTGTGTAATCTTGAAGACATTAAGC 471

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES
source
Location/Qualifiers

1..660
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="5930424019"
/clone_1db="RIKEN full-length enriched, 13 days embryo forelimb"
/sex="mixed"
/tissue_type="forelimb"
/dev_stage="13 days embryo"
/lab_host="DH10B"
/note="Site_1: Salt: Site_2: BamHI: cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGATTAATTAATATCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified p Bluescript KS(+) after bulk excision from lambda FLC I. Cloning sites, 5' end: Salt: 3' end: BamHI."

BASE COUNT 228 a 109 c 150 g 173 t
ORIGIN

Query Match 16.3%; Score 385; DB 9; Length 660;
Best Local Similarity 77.1%; Pred. No. 2.3e-48;
Matches 513; Conservative 0; Mismatches 135; Indels 17; Gaps 3;

QY 1659 gtcaaccacagaagaagtcattcctcgt-ttctcgaaagctgttcggaatcattat 1717
DB 1 gtcacctaaagacaaagtcattctttgttcacagaaaaagccgttcgattcattat 60
QY 1718 caagtacctaactctggaataatcaatagagtccttcgcatgagatagagaacaga 1777
DB 61 caagtattctgattcttcgacacatattcagtgagtccttcgacagcgaacagacaga 120
QY 1778 gaagtcggaagaagcattagagaactttaaacagcaagtgagaactactatgcga 1837
DB 121 gtaaccgagaaagcgttaagaaactttaagacagtaagtgagaaatgatttattgcta 180
QY 1838 ctgactcagcctcctagagagactgagtcacatgacgtttacacatgtctataatttga 1897
DB 181 ctgacttgagcattctgagagcgtttgattgacatgattacatgatttattgattt 240
QY 1898 ttccacgagaatttgaagaatcagtcacacggaatagagcgcaggaagaagcagaagga 1957
DB 241 ttccacgagaactgaaagaaatgacacagagtagagcgcactggagagcagagcgca 300
QY 1958 ctggtgttccacttaacacttgaactagaaatgagtcggaaggttcctcctgaattgata 2017
DB 301 ctggatgtcatttacccttattacaaagaatgattgagagcgttcacactgattgattga 360
QY 2018 atattctgaaagaagaatcagagttatccagaggaagcttgatcaatggctgaagagt 2077
DB 361 atattctgaaagaagaatcagagttatccagaggaagcttgatttattggctgaagagt 420

QY 2078 ttgaagcacatcaacggaagaaggaatggaagaataatggaagaacctcaagaagagc 2137
DB 421 acaaagcaaatataacttaaaagagagatggaataatgggaagacctcaagaaagc 480
QY 2138 ccaagaagcttcaatcaatgctctcgtactagtgaggtagaagaattcaagaattttaga 2197
DB 481 ccgaagaagtattactaatgttgcgtgagacactagatgccgagagattcaagattgtagt 540
QY 2198 aatatagtaagaacagaagatctggacatgctggcagatataagaacccggaactgttga 2257
DB 541 caaagtg-----gagcttgaacatgactgactatgagagac-----ttgg 584
QY 2258 ctgattcttaataatagtgcttgaaataatagaatccagtgcttataacttcttaa 2317
DB 585 ctgattcttaataatagtgcttgaataatgagattcaattgatttatttctttaa 644
QY 2318 taataa 2322
DB 645 taaaa 649

Search completed: July 25, 2002, 11:05:24
Job time: 6999 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 10:20:20 ; Search time 369.6 Seconds

(Without alignments)
10986.208 Million cell updates/sec

Title: US-09-923-831-42

Perfect score: ~ 2365
Sequence: 1 ttgtgaccgagctcgtgctcc.....aaaaaaaaaaaaaaaaaa 2365

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_032802:**
1: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Match Length	DB	ID	Description
1	2365	100.0	2365	20	AAZ32206 Human sdp3.8 (HAGE
2	358	15.1	2210	23	ABL13469 Drosophila melanog
3	307.8	13.0	323	20	AAZ32209 Human sdp3.8 (HAGE
4	253.2	10.7	1098	23	AAZ32200 DNA encoding novel
5	229	9.7	4384	23	ABL13468 Drosophila melanog
6	227.6	9.6	2321	22	ABA82027 wound healing rela
c	201.8	8.5	2888	22	AAZ40962 DNA encoding nove
8	192	8.1	3247	23	AAZ56838 DNA encoding novel
9	186.4	7.9	3825	21	AAZ29604 Human cell cycle r

10	184.8	7.8	3347	22	AAZ68400	Human lung tumour
11	184.8	7.8	3825	20	AAZ91484	Human CTGF30 prote
12	182.2	7.7	771	21	AAZ86782	T. thermophila Hcl
13	174	7.4	1464	23	AAZ54662	Staphylococcus aur
14	174	7.4	1530	23	AAZ51619	Staphylococcus aur
15	174	7.4	3330	23	ABL16829	Drosophila melanog
16	172.8	7.3	2720	23	ABL11075	Drosophila melanog
17	172.8	7.3	4720	23	ABL11074	Drosophila melanog
18	172.8	7.3	4720	23	ABL16822	Drosophila melanog
19	172.4	7.3	1464	23	AAZ55256	Staphylococcus aur
20	172.4	7.3	1677	22	AAZ00826	S. aureus DNA enco
21	170.8	7.2	1584	20	AAZ36519	S. aureus dbpa cod
22	165	7.0	3655	22	AAZ80335	Human immune/haema
23	165	7.0	3655	22	AAZ80335	Human immune/haema
24	164.8	7.0	1575	23	AAZ55756	Human immune/haema
25	164.8	7.0	3230	19	AAZ52336	Streptococcus pneu
26	163.4	7.0	2653	23	ABL16085	Drosophila melanog
27	163.4	6.9	3655	22	AAZ80335	Human immune/haema
28	162.6	6.9	7363	18	AAZ74330	Human immune/haema
29	158.2	6.7	2224	22	AAZ06334	Staphylococcus aur
30	153.2	6.5	2344	21	AAZ36986	Human yasa cDNA
31	153.2	6.5	2348	21	AAZ37726	Arabidopsis thaili
32	151.4	6.4	4209	23	ABL10749	Arabidopsis thaili
33	151	6.4	490	22	AAZ47071	Drosophila melanog
34	151	6.4	490	22	AAZ17641	Human breast cance
35	149.8	6.3	2297	24	AAZ90854	Human breast cance
36	148.6	6.3	1611	23	AAZ53040	Bacillus anthracis
37	148.2	6.3	2725	23	ABL20793	Enterococcus faeca
38	148.2	6.3	2823	23	ABL16329	Drosophila melanog
39	148.2	6.3	2856	23	ABL20791	Drosophila melanog
40	148.2	6.3	2903	23	ABL17833	Drosophila melanog
41	145.4	6.1	2964	23	ABL15506	Drosophila melanog
42	145	6.1	2320	21	AAZ37725	Arabidopsis thaila
43	144.6	6.1	2601	23	ABL16823	Arabidopsis thaila
44	143.4	6.1	4416	19	AAZ69633	Drosophila melanog
45	142	6.0	505	22	AAZ11582	Dead Box Y (DBY) g
						Human brain expres

ALIGNMENTS

RESULT 1	
AAZ32206	
ID	AAZ32206 standard; cDNA; 2365 BP.
AC	AAZ32206;
DP	14-JAN-2000 (first entry)
XX	
DE	Human sdp3.8 (HAGE) encoding cDNA.
XX	
KW	Human: sdp3.10; SAGE: sdp3.8; HAGE: sdp3.5; TRAP: sarcoma;
KW	tumour rejection antigen precursor; tumour associated nucleic acid;
KW	carcinoma; Cancer; Immune response; diagnosis; ss.
XX	
OS	Homo sapiens.
XX	
FB	Key
FP	CDS
FT	Location/Qualifiers
XX	208..2154
XX	/tag= a
XX	/product= "sdp3.8 (HAGE)"
XX	
PN	W09953061-A2.
XX	
PD	21-OCT-1999.
XX	
PF	14-APR-1999; 98MO-US08163.
XX	
PR	15-APR-1998; 98US-0060706.
XX	
PR	27-OCT-1998; 98US-0123989.
XX	
PR	30-OCT-1998; 98US-0185706.
XX	
PR	30-OCT-1998; 98US-0185789.
XX	

PA (LUDW-) LUDWIG INST CANCER RES.

XX Martelange V, De Smet C, Boon-Falleur T;

XX MPI: 1999-620430/53.

DR P-PSDB: AAY49636.

PT New nucleic acid encoding sarcoma-associated gene products, useful for
XX diagnosis, e.g. treating and preventing cancer.

PS Claim 1: Page 88-91; 93pp; English.

XX
CC The present sequence encodes human sdpr.8 (HAGE), a sarcoma-associated
CC gene product (I). Agents, specifically sarcoma associated nucleic acids
CC (II) or their expression products that are tumor rejection antigens
CC (TRA), that selectively increase formation of HLA (human leucocyte
CC antigen)/(I) complexes are used for treating cancer, especially sarcoma
CC and carcinoma, in humans and other animals. Compositions containing
CC autologous cytolytic T cells (CTL), specific for the HLA/(I) complex,
CC are similarly useful, also transformed cells that stimulate such CTL
CC in vivo. (II) are also used: (i) as source of therapeutic antisense
CC sequences that reduce expression of (I); (ii) for recombinant
CC production of (I); (iii) particularly its fragments, as primers and
CC probes in usual hybridisation and amplification assays, for diagnosis,
CC prognosis and monitoring of tumours, or for measuring binding
CC specificity of HLA molecules or CTL clones; (iv) to identify related
CC sequences; and (v) for generating transgenic animals, e.g. for studying
CC cancer and immune responses to it. (I) are used to raise specific
CC antibodies (Ab) and therapeutically. Ab are used to diagnose tumours in
CC immunosays, also for delivering drugs, toxins, imaging agents etc. to
CC (I)-expressing cells.

SO Sequence 2365 BP; 775 A; 420 C; 592 G; 578 T; 0 other;

Query Match 100.0%; Score 2365; DB 20; Length 2365;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgttacccagctccagatccctagtaaacgcccagctgtccgtaagaaagtcggcgga 60
Db 1 ttgttacccagctccagatccctagtaaacgcccagctgtccgtaagaaagtcggcgga 60
Oy 61 tagagagcgtggcgggggggttagcctcgtgggggtcccttaagagtgctgtgctgc 120
Db 61 tagagagcgtggcgggggggttagcctcgtgggggtcccttaagagtgctgtgctgc 120
Oy 121 ttccctgcaacgctagctcttaacgagctacagtggttgcaagagctggagcggcaacg 180
Db 121 ttccctgcaacgctagctcttaacgagctacagtggttgcaagagctggagcggcaacg 180
Oy 181 agctcgagagcgccctctcttgaaacaatgtccacacaggaagagctcccaagcctc 240
Db 181 agctcgagagcgccctctcttgaaacaatgtccacacaggaagagctcccaagcctc 240
Oy 241 acgttgggtcgttcgtacgtcggaagctcgacagtgctccgagcgcagagagggcgcg 300
Db 241 acgttgggtcgttcgtacgtcggaagctcgacagtgctccgagcgcagagagggcgcg 300
Oy 301 gcgagagagtgtaatcgaaacagctcctgagatataatgttcgacagagagtggtcgtcg 360
Db 301 gcgagagagtgtaatcgaaacagctcctgagatataatgttcgacagagagtggtcgtcg 360
Oy 361 agagagacactcagagcccccggagcgctggcgcgtggttcacagagaaactgcgcgtgtt 420
Db 361 agagagacactcagagcccccggagcgctggcgcgtggttcacagagaaactgcgcgtgtt 420
Oy 421 ttgtcttgaagagcactctgttgcgcggttaatcgctcgtgtgttggtcacaataaag 480
Db 421 ttgtcttgaagagcactctgttgcgcggttaatcgctcgtgtgtgttggtcacaataaag 480
Oy 481 aatatacaagtaacaacaacccaatccaatataatataagaacaacacagatcatata 540
Db 481 aatatacaagtaacaacaacccaatccaatataatataagaacaacacagatcatata 540

Db 481 aatatacaagtaacaacaacccaatccaatataatataagaacaacccaatcatata 540
Oy 541 gtcaaaatttttgcaagcaaggaatgcaaaacgaaagcaaaagcagtgatagaacaattt 600
Db 541 gtcaaaatttttgcaagcaaggaatgcaaaacgaaagcaaaagcagtgatagaacaattt 600
Oy 601 gtcaaaagctgaagaaatttacaattcagaaatgtcgaaatgtgatactgattccaact 660
Db 601 gtcaaaagctgaagaaatttacaattcagaaatgtcgaaatgtgatactgattccaact 660
Oy 661 tctgttggaagaagatggaagcacagataacaatgtgtgtgcaagagatcggtcattgata 720
Db 661 tctgttggaagaagatggaagcacagataacaatgtgtgtgcaagagatcggtcattgata 720
Oy 721 gatttggaatcaattagagaaagagtttgaattgcaaaacaaagtcggacagatta 780
Db 721 gatttggaatcaattagagaaagagtttgaattgcaaaacaaagtcggacagatta 780
Oy 781 gatttggaatcaattagagaaagagtttgaattgcaaaacaaagtcggacagatta 780
Db 781 gatttggaatcaattagagaaagagtttgaattgcaaaacaaagtcggacagatta 780
Oy 841 gaagcagatagttggaaggaagaaatttataataacgctggagatgagtggaagtgagg 900
Db 841 gaagcagatagttggaaggaagaaatttataataacgctggagatgagtggaagtgagg 900
Oy 841 gaagcagatagttggaaggaagaaatttataataacgctggagatgagtggaagtgagg 900
Db 841 gaagcagatagttggaaggaagaaatttataataacgctggagatgagtggaagtgagg 900
Oy 901 gagaagagcctatcccaatccatctgacatgtgaaagtcgacagtcgacatgtcattcc 960
Db 901 gagaagagcctatcccaatccatctgacatgtgaaagtcgacagtcgacatgtcattcc 960
Oy 961 gaggatagaaataataaaagcaggttttcaaaagccaacacatttcagtcacag 1020
Db 961 gaggatagaaataataaaagcaggttttcaaaagccaacacatttcagtcacag 1020
Oy 1021 gcatggccatattgttgcgaaggaatagatcttataagagtgagcccgagctggaacaga 1080
Db 1021 gcatggccatattgttgcgaaggaatagatcttataagagtgagcccgagctggaacaga 1080
Oy 1081 aagacatgtgttataatttaatgctgattatcatatcgtgtccttcaacccagcttaaa 1140
Db 1081 aagacatgtgttataatttaatgctgattatcatatcgtgtccttcaacccagcttaaa 1140
Oy 1141 ggtcacaaggaatagacccggcagatgttgaatctcaatcccaatcggaattagcaattcaa 1200
Db 1141 ggtcacaaggaatagacccggcagatgttgaatctcaatcccaatcggaattagcaattcaa 1200
Oy 1201 gtgaagaggaatgttgcacaatatcataaaagagcttcggaagtggttgtgtataatgt 1260
Db 1201 gtgaagaggaatgttgcacaatatcataaaagagcttcggaagtggttgtgtataatgt 1260
Oy 1261 ggtcgaatagagatggaacaatatagaaagagcttaaaaaagtggtatcatataatgtga 1320
Db 1261 ggtcgaatagagatggaacaatatagaaagagcttaaaaaagtggtatcatataatgtga 1320
Oy 1321 actcccggaagatggaatgatacttcgcaaatgaacttcgtcaatctggaagaataaacc 1380
Db 1321 actcccggaagatggaatgatacttcgcaaatgaacttcgtcaatctggaagaataaacc 1380
Oy 1381 tacttgatttataagaaagcagacaagatcttgacatggaatgtgaaacccagataatg 1440
Db 1381 tacttgatttataagaaagcagacaagatcttgacatggaatgtgaaacccagataatg 1440
Oy 1441 aagaatttctagatgctgcccgaagcagcagatgtatgaaacagtgcttaacatgacct 1500
Db 1441 aagaatttctagatgctgcccgaagcagcagatgtatgaaacagtgcttaacatgacct 1500
Oy 1501 catctagttcaacgctcgcgcaaatcttatttgaagaacaaatgattgtctatgttgt 1560
Db 1501 catctagttcaacgctcgcgcaaatcttatttgaagaacaaatgattgtctatgttgt 1560
Oy 1561 acattgatactgagctgtaagtcagtgaaagcaaatataatgtataaccacagagaa 1620
Db 1561 acattgatactgagctgtaagtcagtgaaagcaaatataatgtataatgtataaccacagagaa 1620

Qy	1621	gagaaatgagatgcacatgcaaaactttctacagatgatagctaccacagaaagtcat	1660
Db	1621	gagaaatgagatgcacatgcaaaactttctcagagatgatagctaccacagaaagtcat	1680
Qy	1681	gtcttcgcttcctcgaaagacgctgttcgagatcactatcaagtgcctcaacttggaaat	1740
Db	1681	gtcttcgcttcctcgaaagacgctgttcgagatcactatcaagtgcctcaacttggaaat	1740
Qy	1741	atctcagctgagctctctgctcgtgagatagagaaacagagatcgggagaaagatbaag	1800
Db	1741	atctcagctgagctctctgctcgtgagatagagaaacagagatcgggagaaagatbaag	1800
Qy	1801	aactcttaaaacacgagcaaaagctbaagataatcaattgcgaactgtatcagagcttaagagact	1860
Db	1801	aactcttaaaacacgagcaaaagctbaagataatcaattgcgaactgtatcagagcttaagagact	1860
Qy	1861	gagtgccatgacgcttaacacatgctataatttgacattccacgaaatgtgaagataac	1920
Db	1861	gagtgccatgacgcttaacacatgctataatttgacattccacgaaatgtgaagataac	1920
Qy	1921	gttcacacgagatagggcagacacgagagagacagagagagactgtgttccatccacacttg	1980
Db	1921	gttcacacgagatagggcagacacgagagagacagagagagactgtgttccatccacacttg	1980
Qy	1981	accagaatgacttgagaggtgtgctcctctgaattgataataattctcgaaagagcaaatcag	2040
Db	1981	accagaatgacttgagaggtgtgctcctctgaattgataataattctcgaaagagcaaatcag	2040
Qy	2041	agatattcccgagagagactgctgatactcaatgagcttcgagagaggtttggagacacacgaagaaag	2100
Db	2041	agatattcccgagagagactgctgatactcaatgagcttcgagagaggtttggagacacacgaagaaag	2100
Qy	2101	gaaaatgagaaagaaaatgagaaagacctcaagaaagagcccaagaaagtccaataatgctct	2160
Db	2101	gaaaatgagaaagaaaatgagaaagacctcaagaaagagcccaagaaagtccaataatgctct	2160
Qy	2161	ctgtacatgctgggtgagagaaatccaagatatttttgaagaaatgagtaagacagaaatgtg	2220
Db	2161	ctgtactagctgggtgagagaaatccaagatatttttgaagaaatgagtaagacagaaatgtg	2220
Qy	2221	gaatctgtgcgagctatgaaagacgagactgtatgcactgtactttaaataataatagtgct	2280
Db	2221	gaatctgtgcgagctatgaaagacgagactgtatgcactgtactttaaataataatagtgct	2280
Qy	2281	tgaataatagaatcccgctttctatacttctttaaataaataatagaatgatttaact	2340
Db	2281	tgaataatagaatcccgctgtttataacttctttaaataaataatagaatgatttaact	2340
Qy	2341	aaaaaaaaaaaaaaaaaaaaaa 2365	
Db	2341	aaaaaaaaaaaaaaaaaaaaaa 2365	

RESULT	2
XX	ABLI13469
ID	ABLI13469 standard; cDNA: 2210 BP.
XX	
AC	ABLI13469;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 34889.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical; gene; ss.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	

PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR P-PSDB; ABB69366.
 XX
 PX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 PS Claim 1; SEQ ID NO 34889; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABU01840-ABU16175), expressed DNA
 CC sequences (ABU01840-ABU16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 2210 BP; 598 A; 509 C; 615 G; 488 T; 0 other;

[illegible]

Oy	1292	ttaaanaagvfytagatatcatlaatttgcacaaccccggaagaattgaatgactctgcaatg	1351
Oy	1293	ttaaanaagvfytagatatcatlaatttgcacaaccccggaagaattgaatgactctgcaatg	1352
Db	1306	tgaagcgcgvtgcggaagatcatlctatctgcaactccgggacgcgctgaaacgatttgaataatg	1365
Oy	1352	gtaactctgcaatctggaagatatataacacttctgttttgaatgaagcgcgaagaatgt	1411
Db	1366	ccaatctcatgtatgtaagcaacatcaacttctctgttcttgcagagcgagatcgcatacg	1425
Oy	1412	tgaacatggtatttgaacaccgagataagaatttctgtatgagtgtgcgccagataagc	1471
Db	1426	tgaacacgtgttctcgagccgagacagatccgggaagtgatgctgaataacgtccggaacgtc	1485
Oy	1472	agccagttctgacccgvtgctacataggcccatcttgcattgcacgcctccgcaactctat	1531
Db	1486	agaccataatgacgtccgcccacttgcgcacccagagtagtgcgttttagcgcaagactata	1545
Oy	1532	tgaagaagaacaaatgtattgtctatgtttgtgtaatttgtaattgtctgcgtgaagttaag	1591
Db	1546	tgaagaagaacccctccagagvtgvtgtgtcgatgctgcacgtgcgaacgaacactcgtga	1605
Oy	1592	agdaaaataataatctgtaaccaccgcgagga--agagaaatgtagtgcatactgaacacttcc	1648
Db	1606	aacaaatttlaattgtgtggaatgatagtacatgacgaataataacacatcaactcttcg	1665
Oy	1649	tacaagatgtatgcatccacgagacaagaatgcatgtcttcgttccgaaaagcgtttgcg	1708
Db	1666	ctagaagacgtgccagtaacggaacagatacatlcttctgagcgaagaagttcgtgcgtg	1725
Oy	1709	atcaacttcaagatgaaaccttaataatttggaaataatactagtagagctctctgtaagata	1768
Db	1726	acgaactatccagtgaaactctgcgtgtgattgttcatagaccagatgcatctcatgtatacc	1785
Oy	1769	gagaaacagagagatctcgggaagagaaacattgagaacatttaaaacgggacgaagtggaataac	1828
Db	1786	gcgactaagatgtatcgtgagcagagctatcttgcggaataaagttccggcgctgtgcgcatc	1845
Oy	1829	taatctgaacatgactctagcctctagaagactgtatgtccatgacgttacaacatgtctata	1888
Db	1846	tgttctctacccgattgtgcatacgaatgtgccttggaaattgaaatgaaatatacaacatgtatca	1905
Oy	1889	atttgaacttccaaagaaatcttgaagaataatgtaacacggaataggcgacgaaggaag	1948
Db	1906	acataatatttccgacaacaactcggagagatagtgcacggtgtgtgtgcgaacccggaactg	1965
Oy	1949	cagagagagactgtgttccattacaacacttgaactagaatgatitggaggtgtgcctctg	2008
Db	1966	ctgscggaacagggacacataaataagctttttttagcggagagatttggctatgtgcgaag	2025
Oy	2009	aatttgaatlaattcttggaaaagacaaataacagatctccgagaaagattctgatacatg	2068
Db	2026	aactaatttgaatgtaactgtagaggtggtgcgagcaggaagtgtccgaacgaacttcaaaaatgtg	2085
Oy	2069	ctgtagaggttttaggc 2084	
Db	2086	cttagacgctttaaagc 2101	

RESULT 3

ID AA2

ID AAZ32209 standard; CDNA; 323 BP.

AC AA232209;

DT 14-JAN-2000 (first entry)

Human sdp3.8 (HAGE) nucleic acid fragment.

Human; sdp3.10; SAGE; sdp3.8; HAGE; sdp3.5; TRAP; sarcoma;

KW carcinoma: cancer: immune response: ss,
tumour rejection antigen precursor; tumour associated nucleic acid,

KW carcinoma; cancer; immune response; diagnosis; ss.

OS Homo sapiens.

XX, 1900

Pt	M09J933061-A2.
Pd	21-OCT-1999.
Pf	14-APR-1999:
Px	99W0-US08163.
Px	15-APR-1998:
Pr	98US-0060706.
Pr	27-JUL-1998:
Pr	98US-0122989.
Pr	30-OCT-1998:
Pr	98US-0183706.
Pr	30-OCT-1998:
Pr	98US-0183789.
Pa	(LUDM-) LUDMWG INST CANCER RES.
Pi	Martelange V, De Smet C, Boon-Fallieur T;
Dx	WP1: 1999-620430/53.
Xx	New nucleic acid encoding sarcoma-associated gene products, useful for
Pt	diagnosing, e.g. treating and preventing cancer -
Xx	
Ps	Claim 8: Page 87; 93pp: English.
Cc	The present sequence represents a nucleic acid fragment from human
Cc	sdp3.8 (HAE), a sarcoma-associated gene product (I), which is used to
Cc	derive probes and primers useful in identification. Agents, specificlly,
Cc	sarcoma associated nucleic acids (II) or their expression products that
Cc	are tumour rejection antigens (TRA), that selectively increase formatio
Cc	of HLA (human leucocyte antigen)/(I) complexes are used for treating
Cc	cancer, especially sarcoma and carcinoma, in humans and other animals.
Cc	Compositions containing autologous cytolytic T cells (CTL), specific f
Cc	the HLA/(I) complex, are similarly useful, also transformed cells that
Cc	stimulate such CTL in vivo. (II) are also used: (i) as source of
Cc	therapeutic antitense sequences that reduce expression of (II); (ii) f
Cc	recombinant production of (I); (iii) particularly its fragments, as
Cc	primers and probes in usual hybridisation and amplification assays, fo
Cc	diagnosis, prognosis and monitoring of tumours, or for measuring bindin
Cc	specificity of HLA molecules or CTL clones; (IV) to identify related
Cc	sequences; and (V) for generating transgenic animals, e.g. for study in
Cc	cancer and immune responses to it. (I) are used to raise specific
Cc	antibodies (Ab) and therapeutically. Ab are used to diagnose tumours i
Cc	immunoassays, also for delivering drugs, toxins, imaging agents etc. t
Cc	(I)-expressing cells.
Sq	Sequence 323 BP; 120 A; 55 C; 72 G; 76 T; 0 other:
Qy	Query Match
Bc	Best Local Similarity 13 0%; Score 307.8; DB 20; Length 323;
Mt	Matches 321; Conservative 0; Mismatches 2; Indels 2; Gaps
Oy	727 gatcaaataggagagaagtgttgaaatggcaaaaacaaagtggcgattaccacca 786
Db	1 gatcaaataggagagaagtgttgaaatggcaaaaacaaagtggcgattaccacca 60
Oy	787 attaagaaaaacttttaaaagagtcacctgcgcacaagtggccatgctaaagttagaac 846
Db	61 attaagaaaaacttttaaaagagtcacctgcgcacaagtggccatgctaaagttagaac 120
Oy	847 gatagtctggaggaagaagaatttaataataacgtggatgacttgaagaatgggagaaga 906
Db	121 gatagtctggaggaagaagaatttaataataacgtggatgacttgaagaatgggagaaga 180
Oy	907 cgaaccatcccccatctaccctgacatttgttagcgccttcaattgatccctgaagt 966
Db	181 cgaaccatccc--aatctacctgcacatttgttagcgccttcaatgattaccctgaagt 238
Oy	967 atggaaaacaataaaagagcgagttttcaaaagccaacaactatagctcaagcctg 1026
Db	239 atggaaaacaataaaagcaggtttttcaaaagccaacaactatagctcaagcctg 298
Oy	1027 ccacattgtgttgacaaggaatagac 1051
Db	299 ccacattgtgttgacaaggaatagac 323

RESULT 4

AAST0200
ID AAST0200 standard; cDNA; 1098 BP.
XX
AC AAST0200;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #6004.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-PSDB: ABB06013.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID NO 6004; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (II) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SO Sequence 1098 BP; 279 A; 306 C; 295 G; 218 T; 0 other;

Query Match

Best Local Similarity 10.7%; Score 253.2; DB 23; Length 1098;
Matches 267; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

OY 731 aaattagagaggaagtttgaaatgcaaaaacaagaagtgagagattaccacaatta 790
DB 5 aaatctcaaaagctgctgagaaaacaagaagaaatcttgcagattaccaccaatta 64

OY 791 agaaaactttataaagagtcaccctccacaagtcgcatgccaagaatagaacagata 850
DB 65 agaaaactttataaagagtcaccctccacaagtcgcatgccaagaatagaacagata 124
OY 851 gtctgagaaagaaaatttaataaagctgagctgagctgagagatggagagaacgac 910
DB 125 gtctgagaaagaaaatttaataaagctgagctgagagatggagagaacgac 184
OY 911 ctatcccaatctctaccctgcacattgatgacgcttcaatgtatccctgaagtatgg 970
DB 185 ctatcccaatctctaccctgcacattgatgacgcttcaatgtatccctgaagtatgg 244
OY 971 aaacatlaaaaaggaaggtttccaagaagccaaccttcaagtcacag 1020
DB 245 aaacatlaaaaaggaaggtttccaagaagccaaccttcaagtcacag 294

RESULT 5

ABL13468
ID ABL13468 standard; cDNA; 4384 BP.
XX
AC ABL13468;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 34886.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
DR P-PSDB: ABB69365.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 34886; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SO Sequence 4384 BP; 1235 A; 979 C; 1024 G; 1146 T; 0 other;

Query Match

Best Local Similarity 9.7%; Score 229; DB 23; Length 4384;
Matches 633; Conservative 0; Mismatches 495; Indels 77; Gaps 3;

```

Oy 764 caaagtggcagattaccaccaattagaanaactttatataaagatccatccacaa 823
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1821 acgggtggtcccaagtgccaccgcttaacaaagaacttaccagagctccgaaagt 1880
Oy 824 gtgcacgtcacaag taagaagcagatagttgaggaagaatcttaataac----- 878
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1881 ccaatctgacaaatcagagattgagcgcatccgagagagaacaagaataacagttc 1940
Oy 879 -----gtggatgacttgaaagatggggaagaagaccatcccaatccctcgcga 931
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1941 ctatgttttcgacaaagagggggaacgltcacgcgcctatcccaatccagtttga 2000
Oy 932 cattgtagcgccttcaatgltatccctgaaggtatgaaaaacttaaaaggcaggtc 991
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2001 ccttcgagcaatgcttgcgcaatccagatatgcttgaggaga ttaccaaattggcct 2060
Oy 992 ttcaaaagcccaacacttccagtcacagcagcagcagcagcagcagcagcagcagc 1051
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2061 tcccaagccctcgccatccatccgacgctgagccatctctctccaaagtgatgaca 2120
Oy 1052 ctatgagtagccagacgtggaacaggaagacatgtgttaattaaatgcctgattta 1111
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2121 tgatcgtatccgccaacagcactgtgcaaaactaagcttcccttctgcccgaatga 2180
Oy 1112 ttaactcgtctcctcaacccagccttaaggtcaaaaggaatagaccgcgcatgttattc 1171
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2181 ttccacacgagatccagagcagcagcagcagcagcagcagcagcagcagcagcagc 2240
Oy 1172 taaactccactcgggaattagcacttcaagtagaagagaaatgttgcataatcat--- 1228
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2241 tggcaccactcgggagtgtgcttgcaaaatcgaatlgaggtlcaagaagctactccttc 2300
Oy 1229 -----a 1229
Db 2301 gtgcagataaagcgttaatatccatccgaacatgaaaggttcccttttaacaaaa 2360
Oy 1230 taaaggcttcgagatgtttgtgtatattgtgtgtgaataagatgacaataagaaga 1289
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2361 taaacaaatttgcaaggttctgctgtacgggtggtgcacacacatgacagatccga 2420
Oy 1290 gcttaaaaaggtgtgatatcatatgtcaactccgcggaagatlgaaatgactgcgaat 1349
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2421 ctggagcggtgtgagagatattatctgactccgcggaacagatgttgatga 2480
Oy 1350 gagtaactctgtcaatctgaagaataataactactgtgttagatgaaagcagaagat 1409
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2481 ggcacatgcatctgatatgaaacacatcaactctgtgtgactgaggaagcagatgcac 2540
Oy 1410 gctgacatggaatttgaaacccagataaagaattctgttagatgtgcgccagatag 1469
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2541 gctgacatggttctcgagccgcagatccgaaagtgtatgacatagctgcgcgcg 2600
Oy 1470 gcagacagttatgacagagtcacatgtgccttcaatgcctgcctgcgcgaactctta 1529
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2601 tcagacataatgacgtcggccactgtgcacacaggaagatcgttcttgagcgagagcta 2660
Oy 1530 ttgaaagaaaccaaagtatgtctatgtgtgtacatlggactcgtgtgtcgttaagttcagt 1589
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2661 tatgaagaatccatccacaggtgtgtgtcgatctgcctgacatgagccacgcagctgct 2720
Oy 1590 gaagcaaaataatgttaacacacagga---agaaataatgagacatgcaaaacttt 1646
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2721 gaacaaatataatgtatgagagatgacatgaaatcaaacacatataactttt 2780
Oy 1647 tctacagatagtcatccacagacaaagtcattgtcttccttcgaaagcctgtgtgc 1706
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2781 cgttaagaacatgtccagtaacgacaagatcatatatttctgagcgaagaagtctgcgc 2840
Oy 1707 ggtactcttacaagtacacttaactctgaaataatctagtagaagcttcttcacatggaga 1766
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2841 tgaagacataatccagaaacttaacgtcgtgatttcaatgacccaaatgcatcatcagtgaa 2900
Oy 1767 taagaaacagagatcgggagagaagcattagagaactttaaaacaggaagatgagaat 1826

```

```

Db 2901 tccgatacagatgtagtgcagagctatctgcccagataatgaatccggtgtgcgat 2960
Oy 1827 actaatgcacatgtagctccttagagagacttgatgtccagatgacatgactta 1886
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2961 tctgtgtctccagatgtgcatcaccgctgtgacatttgagatatacagtaagcaaa 3020
Oy 1887 taact 1891
    || |
Db 3021 ttttt 3025

RESULT 6
ABAB2027
ID ABAB2027 standard; DNA: 2321 BP.
XX
AC ABAB2027;
XX
DT 25-JAN-2002 (first entry)
XX
DE Wound healing related polynucleotide SEQ ID NO 118.
XX
KW Human; mouse; vulnerrary; dermatological; skin disorder; wound healing;
XX gene therapy; ds.
XX
OS Unidentified.
XX
PN CA2325226-A1.
XX
PD 17-MAY-2001.
XX
PF 16-NOV-2000; 2000CA-2325226.
XX
PR 17-NOV-1999; 99DE-1055349.
XX 17-DEC-1999; 99US-0172511.
XX 20-JUN-2000; 2000DE-1030149.
XX
PA (SWIT-) SWITCH BIOTECH AG.
XX
PI Regenbogen J, Wolf E, Goppelt A, Werner S, Halle J;
XX
XX WPI: 2001-433142/47.
XX
PT use of novel polypeptide or its variant or nucleic acid encoding the
PT polypeptide for diagnosing and/or preventing and/or treating skin
PT disorders and/or treatment in wound healing or for identifying active
PT substances
XX
PS Disclosure: Page -: 265pp; English.
XX
CC The invention relates to the use of a polypeptide (ABBA44544-ABBA4601,
CC ABBA44606-ABBA4623) or its variant or encoding nucleic acid
CC (ABBA1990-ABBA1995, ABBA2016-ABBA2032) with vulnerrary and/or
CC dermatological activity for the diagnosis, prevention and treatment of
CC skin disorders and treatment in wound healing or for the identification
CC of pharmacologically active substances. The nucleic acids are useful in
CC gene therapy.
CC Note: The printed sequence listing for this specification was incomplete,
CC terminating part way through SEQ ID NO 106. The remaining data was
CC obtained from EPO data for an equivalent patent (EP1114862).
XX
SQ Sequence 2321 BP; 697 A; 458 C; 559 G; 607 T; 0 other;

Query Match 9.6%; Score 227.6; DB 22; Length 2321;
Best Local Similarity 52.9%; Pred. No. 3.3e-40;
Matches 585; Conservative 0; Mismatches 509; Indels 12; Gaps 4;
Oy 950 aagttatccctgaggtttagaaacataaaagcaggtttcaaaagccaaaccta 1009
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 459 acttctcggagatgcatgtagtgtgttgcaaggaagaaacttaaccgaaccacgtcta 518
Oy 1010 ttacgtacagcagatgcccactgtgttgcaaggaatagatcttatagggtgcccaga 1069

```

Db 519 ttcaagctcaggcctggccagctgctcctcagctgagttgatagttcggagtcgcccaga 578
OY 1070 ctggaacagggaagacattgtgtatttaatgctcgtgattatcactcgtccctcaac 1129
Db 579 ctggatctcgggaaaaacattctattctgtcctcgccatctgtacataaacaccagc 638
OY 1130 ccagccttaagggtcaaggatagaaccgcatgttgaattctaaccctccagcgat 1189
Db 639 cactcct--agagagaggtgagtcggccatttgccttgccttggtgcgcacccaggaac 695
OY 1190 tagcacttcaagtcgaaggaggaatgttcgaataatt---cattaaaggcttcggagt 1246
Db 696 tggcaccagcaggctgcagcaagctggtcgtgaattatgtcgaagctgttcgcttgaagtc 755
OY 1247 tctgtctatagctggtggtggaatagaagatgaacaaatagaagagcttaaaaaagctgag 1306
Db 756 ctgtgactctatgctgctgctcccaaggaccacagatctcgtgatttggaaaaggtgtg 815
OY 1307 atatcataattgcaactcccggaagattgaatgactctgcaaaagagtaacttcgtcaacc 1366
Db 816 aatctgtatctgcaacaccctggaagagctgacttcttgaagctggtggaaaaccacatc 875
OY 1367 tgaagataataacccttctgtttagaagcagacagaatgttgcacatggtttg 1426
Db 876 tgaagaagaacaacttacccttgccttgcagagcgtgataagatgcttgatagtggtttg 935
OY 1427 aacccagataatgaagatttctgtagatgtgcgcccagagtcagacagcttatgacca 1486
Db 936 aacccagataatgaagaaatctgtagatcaataagaaccggaatagaagcaacatactgtag 995
OY 1487 gtgtacatagtgcctcattcagttcactgcctgcgcacaacttctatctgaagaacatga 1546
Db 996 gtgcacacttggccaaagaagtaagacagctcgtcgaagattctcctgaagaactatactc 1055
OY 1547 tctgtctatgtgtgacatgtgacatgctgtcgtgaagtcgtgaagcgaataatatactg 1606
Db 1056 atatacatatctggtgcactgtgaactgtgacgtcaaacataaactctcagaattgvgagtg 1115
OY 1607 taacacccga---ggaagagaatcggagtcacatgcacaacttctcaagagtagtcat 1663
Db 1116 tagtcatgagctgcaaaaagatgaaaagcttatctgtcgtatggaagaatcatatgagtg 1175
OY 1664 ccacag---acaaagtcattgtcttcttcgaaaagcgtgtgcgatacacttataca 1720
Db 1176 agaaagaggaataaacaactatgtcttctgtgaacaccaaagaagatgtgatactatacca 1235
OY 1721 gtgacctaatcttgaaataatacagtagagctctcgtacatggaagatagaagaagag 1780
Db 1236 gaaaaatgagagagatggtgtggtcgtccatggtcatcgtgtgacaaagatcagcag 1295
OY 1781 atcggagagaagcatalagagaactttaaaacagcgaagctgagatactaatgcaactg 1840
Db 1296 aacgtgacttgggttcttaaatgaaatccaacatggaagaagccctcattcgtatgctacag 1355
OY 1841 atctagcctcgaaggagctgtgactgacatgacgttatacacaagcttataatttgacttc 1900
Db 1356 atgtggtcctccaagggctagatgtggaagatgtgaaattgtcatcaattatgactcacc 1415
OY 1901 cagcgaatatcgaagaatagcaccaggaatagcgcagcgggaagagcagggagagactg 1960
Db 1416 ctaactcctcagagattatcatcatcgaaatgtggaagaagcgtcgcagtgacaaaagc 1475
OY 1961 ggtttccatatacaacttgcacagaatgtatggaggtgtgcctcgtaatgataata 2020
Db 1476 gcacagcatataacttctcttaacaccttaaaaataaagaagaagtgaagcacttatctg 1535
OY 2021 ttcttggaaaagcacaatcagagtaatt 2046
Db 1536 tgcctcggaagcctaactcaagcaatt 1561

RESULT 7

AAS40962/c
ID AAS40962 standard; cDNA; 2888 bp.
XX
AC AAS40962;
XX
DT 17-DEC-2001 (first entry)
XX
DE cDNA encoding novel human enzyme polypeptide #178.
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW 1,4-lase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticonagulant; ss.
OS Homo sapiens.
PN W020015301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01239.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184684.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0186874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214866.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216847.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0228343.
PR 01-SEP-2000; 2000US-0228344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.

Db 2278 TTACTAAGAGAGTAAAGAGTTTCCAGAGCTTTGGAGCTTAGAGTGCTGTGTTTAGG 2219
Qy 1259 gfygfygaataagagatgaaacaataagaagacttaaaaaagtgatatacaaatg 1318
Db 2218 GAGGAACAGAAATCAGTACAGATGTTGCTGAGCTAAGAAAGAGTGAATTAATGTTT 2159
Qy 1319 caactccggagaatltgaatgactgcaatgaagaaac-----tcgtcaatctga 1369
Db 2158 GCACACCTGCTGGAATGATGACATGTAGCCGTTAACAGTGTGGGTCACAAATCTTC 2095
Qy 1370 agaataaactactgtgttttagaagagcagaaagatcttgaaatgagtttgaac 1425
Db 2098 GAAGAGTACATATGTTGTTTATGATGAAGACAGAAATTTTGACATGCGTTTGAAC 2039
Qy 1430 cccagataaagaatltgttagatgtgcgccagataaggcagacagatgaaccagtg 1489
Db 2038 CCCAGGTCATGCGCATGCTGATTAATGTTGCTGATGACAGACAGGTTATGTTTCAAG 1979
Qy 1490 ctacatggccctcatcagctcagctgcacaaacttatttgaagaaccaatgatg 1549
Db 1978 CTACTTTCCAGAGCTATGAGAGCTTTGGCTGCGAGATCCTTCAAGTAACGTAATGAAG 1919
Qy 1550 tctatgtgtgacatltggatcagttgtcgtgaagtcagtgaaacaaatataatgttaa 1609
Db 1918 TACAGTTGGAGGAGGAGAGTGTGTTTGCTCAGAT---GTGAGACAAACAGTATGTTGA 1862
Qy 1610 ccaacgaggaagaagaatgtgacatgcaaaacttctcagaagatgcatccacag 1665
Db 1861 TTGAGAGAAAGAAAGAAATTTTMAAATTTACTTGAGCTTCTGAGCCATTAACAGAGCTCAG 1802
Qy 1670 acaaaagcatgtcttcgtctcgcgaagaagctgttcgagtaactatacaagtgaactaa 1729
Db 1801 GATCTGCTATTATTTTGTGATTAACAGAGAACATGCTGATGCTTCTTAAGGATTTAA 1742
Qy 1730 tacttgaaataatatacagtagagctctcagatgagatagaagaacagagatcgagaa 1789
Db 1741 TGAGAGCATCTTATCTTGGATGTCTTCAAGGAGCATTTGATTAAGACAGAGATA 1682
Qy 1790 aagcatagagaacttaaaacagcgaagtgagaataactaatgcaactgactagcct 1849
Db 1681 GCATCATTAATGACTTTTAAAGAAATGGAGCTGCAGAACTTGTGCTACCTCTGTTGCTG 1622
Qy 1850 ctagaagactgactgcatcagatgacgttacacatgctataaatttgacttccagagaata 1909
Db 1621 CCCGAGGCTAGATGTGAACATCTGATCTGTAGTAATTAATTAAGTCCCAACACAT 1562
Qy 1910 ttgaagaatactgtaaccgaataagggcgcagggagaagaagcagagagactggttcca 1969
Db 1561 ATGAGGATTTATGTACACAGAGCGGCGGACTGGAAGAGCAAGCAAGGCTTATGCTT 1502
Qy 1970 ttacaacttgactagaagaatgtgaggtgtgctctcgtgaattgattaatctgagaa 2028
Db 1501 ATACTTTTATCAGAGAGATCAAGCTCGCTATGCTGAGCATTAATTAAGGCTTTGAA 1443

RESULT 8
ID AS68838 standard, cDNA, 3247 BP.
XX
AC AS68838;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4642.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX

PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001MO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSQ INC.
XX
PI Dmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
XX
P-PSDB: ABG04651.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 4642; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (II) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 3247 BP; 1089 A; 551 C; 870 G; 737 T; 0 other;

Query Match 8.1%; Score 192; DB 23; Length 3247;
Best Local Similarity 51.3%; Pred. No. 2; 2e-32;
Matches 554; Conservative 0; Mismatches 510; Indels 16; Gaps

Qy 962 aggtatggaaacaacttaaaagcaggtttcaaaaagccacacctttagcagag 1021
Db 1289 agacttaaatccctcaagaagacatgcatgaaagaagccagccctcaaaaccaag 1348
Qy 1022 catggcccatgtgtgcaaggaatagatcctatagaagtagccagactggaacagaa 1081
Db 1349 ctatccctgctataatgcttgacagatattgtgcatctgcaaaacagaagtgaag 1408
Qy 1082 agaatgtgttatttaagctgtgatttcatctgttccttaaccagccttaag 1141
Db 1409 agaacatgtcttctgttgcctcatgtttgaca---catcatgatatcagaagtcattag 1465
Qy 1142 gtcaaggaatagaccggcagctgtagttcctaaaccacactcgggaattagcaattcaag 1201
Db 1466 aggaagagagaggccataagctgtcatcatgactcaactcgaactcgtcttaacaga 1525
Qy 1202 tagaagagaatgtgcaaatcatca---taaaaggtctcgaggtgttggtatag 1258
Db 1526 ttactaaagagtgaagaagtttccaagccttgaggacttgagtggtcgtgttaag 1585
Qy 1259 gfygfygaataagaga tgaacaataagaagacttaaaaaggtgtagatacatcaatg 1318
Db 1586 gaagaaacagaatcagtgagcagatltgtagctgaaagaggtgcgtgaatatltgtt 1645

QY	1205	aagagaaatggtgcgaataattccta---taagggcttcggagtggttctgtatagtg	1261
Db	1163	atccagaatagttaagcgggttttggaaaagcatalaactctcgatcagtgccgtatagag	1222
QY	1262	gtggaatatgagatgaaacaaataaagaaagcttcaaaaagagtgtagatcatatagtcga	1321
Db	1223	gaaggagatgtgagggaagccaaagcccttcagagaggcgagagattgtgtgtga	1282
QY	1322	ctcccgaaagatgtaatgacgtccgaatgaaatgtaacttcgtcaatcctggaataataact	1381
Db	1283	ccccaggtccgacgtatagatcatatgtaaaaagaagctaccatacttcaagaagtcctc	1342
QY	1382	acttggtttttagatgaagcagaagaatgttggacaatggatttgaaccccgataatga	1441
Db	1343	accttgctttttagaagcagatctcgatgttttgaatgttgatttgataccaaatgtcat	1402
QY	1442	agatttgttagatgtggcccgagatagagcaagaagttatgacaagtgtaactgtgctc	1501
Db	1403	cccttagcaagatcatgtctgcgtccgacagcgaaacctcttatattagtgcaccttttgg	1462
QY	1502	atccagttcatatgccttcgcacacatacttatitgaagaaccaagatgttcatgttgta	1561
Db	1463	agaaagaattgaaaagatttggccagaagaaatccttatatcccttatcgaagtgtgacag	1522
QY	1562	catgtgattcatgtctgtcgtatgaatctcagtgaaacaaatataattgtataccaccgag	1621
Db	1523	ataatggagagccaatggaatgtgacacagatgttgtagatttccattccttcgaacta	1582
QY	1622	agaaatagaggtcacatgcaaaaactttctcaagaatgatagtcatccacagaaaagtcattg	1681
Db	1583	gtcaaatggaactgtgcttaaccgggcgtctgttagaatcttaaccttcagaggagtctccc	1642
QY	1682	tctctgttcttcgaaaagctgtgttgcgatacatttcaagtgacctatacttggaaaata	1741
Db	1643	tctttgtctataaaaagcccaatgctggaagactagcagaataaacttaaacagagaggctc	1702
QY	1742	tatcagtgaaagctctgtcatatgagaataagaagaagatcaggagagaagaacattagaga	1801
Db	1703	ataactcttggtcgtctccatagtcgataatgatacgaagtcgaaagaaataagatcattcag	1762
QY	1802	actttaaacaagcagaagtgtgaaataactaatgtgacaactgatacctctagaagacttg	1861
Db	1763	acttttaagaanaaagaaacatcccaagctctgttggccacagaatgttgcagcccggtgtcgg	1822
QY	1862	atgtccatagcgttacaacatgtcatataattgtacttccacgggaatatgaaagaatcgg	1921
Db	1823	acattctctcaattaaagctgttcaatcaatgattgtgcacgggaatttgataagcaca	1882
QY	1922	tacaacgaatagcgcgcacgcgggaagagcagaagactgtgttccattacaacttga	1981
Db	1883	ctcacagagatttggccgcacagaagaagcgggtggaagaagtgtagctataactacta	1942
QY	1982	ctgaagaatgatctggagggatctccctccataattgatctaattctgttgaagaagcaca	2041
Db	1943	ctcccaagacagcaaatatttgcgtgacotgtccgaaacttggaaaggcccaatcaac	2002
QY	2042	gtattccagagagactgtatcaaatgac	2069
Db	2003	acgttctcaagaacactctatagctggc	2030
RESULT 12			
AA288782			
ID AA288782 standard; CDNA: 771 BP.			
AC AA288782;			
DT 18-MAY-2000 (first entry)			
XX T. thermophila HcI cDNA.			
DE			
KW HcI; DEAD box protein; RNA helicase; cytosolic; cancer; diabetes;			

KW	Immunoprotective; anti-diabetic; gene therapy; treatment; diagnosis;
KM	autoimmune disease; multiple sclerosis; rheumatoid arthritis; allergy;
KM	Alzheimer's disease; neurodermatitis; atherosclerosis; osteoporosis;
KM	infection; immunosuppression; ss.
XX	
OS	Tetrahymena thermophila.
PN	WO200005388-A1.
XX	
PD	03-FEB-2000.
XX	
PF	10-JUL-1999; 99WO-EP04892.
XX	
PR	22-JUL-1998; 98DE-1032783.
XX	
PA	(AVERI) AVENTIS RES & TECHNOLOGIES GMBH & CO KG.
PI	Bohmet K, Huels C, Mueliner S;
XX	
DR	WPI: 2000-182709/16.
DR	P-SDB: AAV51577.
PT	New Tetrahymena thermophila genes encoding RNA helicases, useful for
PT	treatment of, e.g. cancer and autoimmune disease -
XX	
PS	Claim 1: Page 42: 54pp: German.
XX	
CC	This invention describes novel Tetrahymena thermophila genes encoding
CC	RNA helicases (eBead box proteins). The products of the invention have
CC	cytostatic, immunoprotective and anti-diabetic activity and can be used
CC	for gene therapy. The nucleic acids or RNA helicase polypeptides can be
CC	formulated with a pharmaceutically acceptable agent and/or adjuvant for
CC	treatment or diagnosis of cancer, autoimmune disease, in particular
CC	multiple sclerosis or rheumatoid arthritis, Alzheimer's Disease, allergy,
CC	in particular neurodermatitis, Type I or IV allergies, arthritis,
CC	atherosclerosis, osteoporosis, acute and chronic infections and/or
CC	diabetes and/or to influence cell metabolism, in particular
CC	immunosuppression for all transplantation's and/or genetic diseases in
CC	particular Werner Syndrome, Bloom Syndrome, xeroderma pigmentosa and
CC	connective tissue disorders. The nucleic acids can also be used to
CC	identify functional interactive molecules or to find variants of RNA
CC	helicases. The nucleic acids or proteins can be used to influence
CC	protein biosynthesis, to inhibit or stimulate degradation of mRNA and/or
CC	stabilize mRNA. The nucleic acids or polypeptides can also be used as
CC	selection markers in molecular biology. The nucleic acid and polypeptide
CC	can be heterologously expressed in useful plants. This sequence encodes
CC	the Tetrahymena thermophila RNA helicase Hcl protein, described in the
CC	method of the invention.
XX	
SQ	Sequence 771 BP: 268 A; 113 C; 153 G; 237 T; 0 other;
QY	
Query Match	7.7%; Score 182.2; DB 21; Length 771;
Best Local Similarity	54.1%; Pred. No. 2,1e-30;
Matches 417: Conservative	0; Mismatches 348; Indels 6; Gaps 2;
QY	1177 ccacatcggaattagcaacttaagtagagagaagatgtgcgaat--atcatataaa 1233
Db	1 ccacatcagagaacctcgctatgcgaatacgsaagaagagcgcgaagaattggttaactctc 60
QY	1234 gggcttcggaagtgttctgtatattggttggtggaattagatagataaataagaagact 1293
Db	61 aagcttaaatgtgcttctatctatgttgytgcctacaaataacccctaaagagcactctc 120
QY	1294 aaaaaagggtagatatacaataatgcgaactcgcgaagatgtgatatcgcgaatgag 1353
Db	121 caataagggtagatgtagtattgtctactctcgttagactatttgactttttagaagc 180
QY	1354 aacttcgcgaatcgsaagaatataaacttaacttggttttagaagaagcagaagaagtct 1413
Db	181 gaaactactacttaagctagatgagttacttatacttcgattatagaaagcagatagaatt 240
QY	1414 gaacttgagtttgaaccacagataatgaagaatttctttaaatttgccacgaattaggca 1473

Dh	241	gataatggtgttttggaaatttaattagaanaaaacttgggttaattagaaactgttcgtttaa	300
Oy	1474	acagttatgcacagttgctacatggtcctacatcagttacatgcctgcgcacaaatcatt--	1531
Dh	301	acattatgtttctctgcctacccgtccgaagaatggtttagaactctgtcttaagaattatgc	360
Oy	1532	-tgaaagaaacacaaatgattgtctatggtttatggtttacatggatcgaattgtctgaattcagtg	1590
Dh	361	aaagaataaccocggtttatglttcaaatctcgaaacaatgaattatgattatlaacgaagaat	420
Oy	1591	aagcaaatataatattgttaaccaaccgggagaagaaatggatcacaatgcacatttcta	1650
Dh	421	aaataataattgttattgtttacatgcataatcaagaataatcaatacttaatacaagaacatla	480
Oy	1651	cagaagatgtatccacacgaacaaagcatatgtcttcgtttctcgaagaagctgttgcgagt	1710
Dh	481	gatgtgttgacttagaagaagataaagattgatatttcgcttaacaaagaagagattgtaa	540
Oy	1711	caattatcaatgtgacctataactttgaaatataatcaatgaattgtctctgatggagataga	1770
Dh	541	agcatatgctgatttttgaataaagaagaagatttaagtgtcttggtatccatgtgtgacaaa	600
Oy	1771	gaacagagatcgcggagagaagaactagaagaactttaaacaaggcaagtgagaatacta	1830
Dh	601	gctataaagaacagagactatgattatgaacaaagttcaaaagcggagatgcgaatacct	660
Oy	1831	attgcacactgactctgcctctagaaggaacttgattgtccattgacgtttacacattgtcataat	1890
Dh	661	attgtctacagaactgatacgaatgaagaagttcgtgaattgaagaatgctctcccgatttaatt	720
Oy	1891	tttgactttccacggaaatattgaataacgcacacccgaataagagcgacag	1941
Dh	721	tacgatttcccaaggttaatgagaagactatgtccacagaatcggtagaagc	771
RESULT 13			
AAS34662			
ID	AAS54662 standard; DNA; 1464 BP.		
XX	AAS54662;		
AC			
XX			
DT	13-FEB-2002 (first entry)		
DE	Staphylococcus aureus DNA for cellular proliferation protein #974.		
XX			
KW	Antisense; ds; prokaryotic cellular proliferation gene;		
KM	antibiotic; antibacterial; drug design.		
XX			
OS	Staphylococcus aureus.		
XX			
PN	WO200170955-A2.		
XX			
PD	27-SEP-2001.		
XX			
FE	21-MAR-2001; 2001WO-US093180.		
XX			
PR	21-MAR-2000; 2000US-191078P.		
PR	23-MAY-2000; 2000US-206848P.		
PR	26-MAY-2000; 2000US-207727P.		
PR	23-OCT-2000; 2000US-242578P.		
PR	27-NOV-2000; 2000US-253628P.		
PR	22-DEC-2000; 2000US-257931P.		
PR	16-FEB-2001; 2001US-269308P.		
XX			
PA	(ELIT-) ELITRA PHARM INC.		
XX			
PI	Hasselbeck R, Ohlsen KL, zyskind JW, Wall D, Trawick JD, Carr GJ;		
PI	Yamamoto RT, Xu HH;		
XX			
XX	WPI; 2001-611495/70.		
PR			
PR	P-PSDB; AAU36803.		
XX			

Pt	New polynucleotides for the identification and development of
Ft	antibiotics, comprise sequences of antisense nucleic acids -
Xx	
Ps	Claim 27: Seq ID No 8299; 511pp; English.
Cc	The invention relates to antisense inhibitors of genes essential to
Cc	prokaryotic cellular proliferation, their use in identifying the
Cc	genes, their use in the discovery of novel antibiotics, the essential
Cc	genes themselves and the encoded proteins. The prokaryotes used are
Cc	<i>Escherichia coli</i> , <i>Staphylococcus aureus</i> , <i>Salmonella typhi</i> , <i>Klebsiella</i>
Cc	<i>pneumoniae</i> , <i>Pseudomonas aeruginosa</i> and <i>Enterococcus faecalis</i> . The
Cc	invention is also useful for the identification of potential new targets
Cc	for antibiotic development. The antisense nucleic acids can also be used
Cc	to identify proteins used in proliferation, to express these proteins,
Cc	and to obtain antibodies capable of binding to the expressed proteins.
Cc	The proteins can be used to screen compounds in rational drug discovery
Cc	programmes. The antisense nucleic acid sequence is also useful to screen
Cc	for homologous nucleic acids which are required for cell proliferation in
Cc	a wide variety of organisms. The present sequence encodes an
Cc	essential prokaryotic cellular proliferation protein.
Cc	Note: The sequence data for this patent did not form part
Cc	of the printed specification, but was obtained in electronic
Cc	format directly from WIPO at
Cc	ftp.wipo.int/pub/published_pct_sequences.
Xx	
S0	Sequence 1464 BP: 540 A; 236 C; 300 G; 388 T; 0 other;
Oy	Query Match 7.4%; Score 174; DB 23; Length 1464;
Oy	Best Local Similarity 52.2%; Pred. NO. 1.6e-28;
Matches 527; Conservative 0; Mismatches 450; Indels 33; Gaps 5	
Oy	988 gttttcaaaagccaacatttcgttaacgtagcgccatgttgttggaagaata 1047
Db	4 ggatttaaaggccgacactatcccanaaagacagtatccctatgcgtacaagaatt 63
Oy	1048 gatccthatagagtagcccgactgtagaacaggaagacatgtygtattattaacgtcggta 1107
Db	64 gatatactcttgycgaagtcacaacgcgtacagtgtaaaccgaggaatcgttatctcta 123
Oy	1108 ttatattcatcttggtccctcaaccgaccttaaggtcaaagatatgaccgccgatgta 1167
Db	124 attggagaagtagtagtaggaacaaagggtlcaatc-----gltg 162
Oy	1168 gtcttaactccccactcgggaattagcaactccaagttagaaggaggaatgttgcaaatatca 1227
Db	163 attcttagccctacaaagaatgtggaacgacggttagcttaaacatatgaagaattcgc 222
Oy	1228 tata---aagggtctcgagtgcttgtagatctgtgttgtaabtagaataagaaata 1284
Db	223 cgtggaacaaagtgfccaaagtggttaactgattcgttggtgtgtagtgcctbcgaagccaattc 282
Oy	1285 gaagagcttaaaaaaggtctagatatcatbaatctgcaacctccggaaagattgaatgctgy 1344
Db	283 aaagccttgaaaanaaaggccacacaaatcgttagtcgnaaaccttgcgtgttacgcacat 342
Oy	1345 caaatgagtaacttcgcgaactcgaagaaataaacacactgtgttttagatgaagacagac 1404
Db	343 ttaaatcgttcgacatttaaaaaaggcgggaattcttaacttgatttagatgaagacgtgat 402
Oy	1405 aagaatgttggaatcgggaatttgtaacccagataatagaagatttctttagatgtgcgccca 1464
Db	403 gaatatgtagtaatactggaatctcagtagatgaagaatttatitagtgaataaalctccaaga 462
Oy	1465 g--atbgacgacagatcatgaacaaagtgaataatgacgtccatcaagtcatcgctcgca 1521
Db	463 gttcaacgtcaacaacatggtgtctcagtactgaataagtcgaataagcaatccaagcttagta 522
Oy	1522 caactctatttggaagaaccaatgatgtctatctgtgtgtagacattgtagcttagtgcgtga 1581
Db	523 caacaactattatgaataatccaaaataatcatttaagaacaatgaataatgaatgtcgtcca 582
Oy	1582 agttcaatgaagcaaaatatatatttgtaaccacccaggaataaagaatgtagtcaatgcaa 1641

Db 583 caaatcgaaagatcctacacattgt---taagaattagaagaatttgatcacattaca 639
Qy 1642 actttctacagagatgcatccacagacaagtcattgtctctgttctccgaaagct 1701
Db 640 aattcctaga---tgctatcaacctgaattgcaatcgatcttgagacgtacnaaacgt 696
Qy 1702 gtcgcgatactacttcaagtgacctaatacttggaaatatacagtagagctctgcac 1761
Db 697 cgtgtgatgaattacaacagtgcttgcatttcttaagaagataaagctgaaggtctacat 756
Qy 1762 ggaagatagaagaagagatcgaggagaacattagaagaacttgaacaagcaagtg 1821
Db 757 ggtgatatcaacaagcgaacagctttagaagatataagaatataaattacacaaatc 816
Qy 1822 agaatactaattgcaactgatactagcctctagaaggacttgatgcatgaagctacacat 1881
Db 817 aatatctagtcgactactgattgtagcagcaagagactagatatcttctgtgtgagtcac 876
Qy 1882 gtcataatttgactctccacggaatataagaagatagctacacggaatagagcgacg 1941
Db 877 gctataacttgatactaccataagatctgaagcctacacacgctattgtctgctacg 936
Qy 1942 ggaagagcagaggagagactggttcttcatacaacttgaactagaatga 1991
Db 937 gctcgtcgtgtaagaaggtatcgcggtacacattgttaactcaattga 986

RESULT 14
AAS51619
ID AAS51619 standard; DNA; 1530 BP.
XX
AC AAS51619;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #36.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW Antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001MO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI: 2001-611495/70.
DR P-PSDB; AAV35760.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Claim 27; Seq ID No 4201; 511pp; English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/publ/ished_pct_sequences.
XX
SQ Sequence 1530 BP; 571 A; 237 C; 309 G; 413 T; 0 other;

Query Match 7.4%; Score 174; DB 23; Length 1530;
Best Local Similarity 52.2%; Pred. No. 1.6e-28;
Matches 527; Conservative 0; Mismatches 450; Indels 33; Gaps 5;

Qy 988 gttttcaaaagccacaacactatcagtcacagggccatctgtgttcgaagaata 1047
Db 121 ggatttaagaagcgcacacactatccaaaagacagatcccttactgcttcagaagaatt 180
Qy 1048 gatcttataggagtagccagactggaacagagaagacatggtttatctaagtcctgga 1107
Db 181 gatatccctgggcaagctccaacggtacacggtacaaacagagagcattccggtacttcta 240
Qy 1108 ttatctatcgtgctcttcaaacccagccttaaaagtcacaaagatagaccgcgcatgta 1167
Db 241 atcgagaagaagtagtgggaacaaagggttcaatc-----gtcg 279
Qy 1168 gttctaacccaccccggaattagcactcaagtgaagagagatggtgcaatatacca 1227
Db 280 attttgacccctacagaagaatttgcagatcaggtatgcgcaacataagaagattagc 339
Qy 1228 tata---aaggcttcgagagtgctgtgatatggtgtggaatagatgacaaata 1284
Db 340 cgtggacaagagtgctccaagtgctactgctatcgtgtggtatgctctacgaagcacaatc 399
Qy 1285 gaagagcttaaaaagagtgatagatcatcaatgcaccccggaagatgaaatgacatcg 1344
Db 400 aagagccttgaaaaaagcccaacatcgtatgctggaacccctggcgctgtatctgacacat 459
Qy 1345 caaatgagtaactcgtcaatcctgagaatataactacttggttttagatgaagcagac 1404
Db 460 ttaaatcgtgcacattaaaacgagcgaattcacaacttgattttagagaagctgcat 519
Qy 1405 aagatgttggacatggaatttgaacccagataaagaatttggatgtgctgcacca 1464
Db 520 gaattgatgatatggaattcatcogatgatatgagatttatatgataaataatccagca 579
Qy 1465 g---atagcagacagttatgacacagtgctacatggtcctcatcagttcatcgctctgca 1521
Db 580 gtacaagctacaacatgttcttcacagctacaatcgtcttaagaacatccaaagctttagta 639
Qy 1522 caatcttatggaagaacaaatgattgtgtctatggtgtgatacttgatcgtatcgctgta 1581
Db 640 caacaatttatgaaatcaccaaaaatcattataagaacaatgataaagaatgtctgaccca 699
Qy 1582 agttcagtggaagcaaatataattgttaaccacgaggaagaagaatgagtcacatgcaa 1641
Db 700 caaatcgaaagatctcatatacttgt---taagaattagaagaatttgatatactaca 756
Qy 1642 actttctacagagatgcatccacagacaagatctgtctctgttctccgaaagct 1701
Db 757 aattcctaga---tgctatcaacctgaattgcaatcgatcttgagacgtacnaaacgt 813
Qy 1702 gtcgcgatactacttcaagtgacctaatacttggaaatatacagtagagctctgcac 1761

Db	814	cgctgctgagatataacacagtcgtcttgattttcttccaagagatataaacgctgaagccttaac	873
Oy	1762	ggagatctgagacagagagatcgggagaaagcattctagagaactttaaaacggcaagtg	1821
Db	874	ggtgatactacacaaacggaaaacggttctagaagctatctaaagaattctaaataacgacaacat	933
Oy	1822	agaaatactaatctgcaactgatacctctctagaagagactgatacgtatccatgacgtataacat	1881
Db	934	aatactttagtcgctactgactgactgacgacgaagagactagatttctcgtgtgagatcat	993
Oy	1882	gtctataatttactctcccggaataatttgaaagaatctacatccgaataaggcgcaag	1941
Db	994	gttataactcttgatactaccacagaatactcgaagaactatacacaccggtatctgctgacg	1053
Oy	1942	ggaaagcagcaggcgactgctgcttctcatatacaactctgacataaata	1991
Db	1054	gtctgctgctgtaagaagagatcgcgcgtaacatttctaataatcgaatga	1103
RESULT 15			
ABL16829			
ID	ABL16829	strand: DNA;	3330 bp.
XX			
AC	ABL16829;		
DT	26-MAR-2002	(first entry)	
XX			
DE	Drosophila melanogaster genomic polynucleotide seq	ID NO 1960.	
XX			
KW	Drosophila; developmental biology; cell signalling; insecticide;		
KW	pharmaceutical; gene; ds.		
XX			
OS	Drosophila melanogaster.		
PN	MO200171042-A2.		
PD	27-SEP-2001.		
PF	23-MAR-2001; 2001WO-US09231.		
XX			
PR	23-MAR-2000; 2000US-191637P.		
PR	11-JUL-2000; 2000US-0614150.		
XX			
PA	(PEXE) PE CORP NY.		
XX			
PI	Venter JC, Adams M, Li PWD, Myers EW;		
DR	WPI: 2001-656860/75.		
XX			
PT	New isolated nucleic acid detection reagent for detecting 1000 or more		
PT	genes from Drosophila and for elucidating cell signalling and cell-cell		
XX	interactions -		
PS	Claim 1: SEQ ID NO 1960; 21pp + Sequence Listing; English.		
XX			
CC	The invention relates to an isolated nucleic acid detection reagent		
CC	capable of detecting 1000 or more genes from Drosophila. The invention is		
CC	useful in developmental biology and in elucidating cell signalling and		
CC	cell-cell interactions in higher eukaryotes for the development of		
CC	insecticides, therapeutics and pharmaceutical drugs. The invention		
CC	discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins		
CC	(AB57737-AB922072).		
CC	The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pcl_sequences.		
XX			
XX	Sequence 3330 BP: 855 A; 923 C; 926 G; 626 T; 0 other:		

Query Match	7.48;	Score 174;	DB 23;	Length 3330;
Best Local Similarity	49.48;	Pred. No. 2e-28;		

[illegible]

Thu Jul 25 15:15:57 2002

us-09-923-831-42.rng

Page 17

Oy	2009	aattgcttaataatctcttgaaagacaaatcaagagatltccaagagagccttgcataatg	2068
Db	2306	aactgactctctgctcgtcgaagagcctcgccaaagccgctgcagaagcctcttctggaatctg	2365
Oy	2069	ct 2070	
Db	2366	ct 2367	

Search completed: July 25, 2002, 12:25:05
Job time: 7485 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 09:11:00 ; Search time 4248.01 seconds
(without alignments)
11650.457 Million cell updates/sec

File: US-09-923-831-42

Perfect score: 2365

Sequence: 1 ttggtaccgagctcgatcc.....aaaaaaaaaaaaaaaaaa 2365

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Minimum number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

- 1: GenBank.*
- 2: gb_da.*
- 3: gb_hlg.*
- 4: gb_in.*
- 5: gb_om.*
- 6: gb_ov.*
- 7: gb_pat.*
- 8: gb_ph.*
- 9: gb_pl.*
- 10: gb_pr.*
- 11: gb_ro.*
- 12: gb_sts.*
- 13: gb_sy.*
- 14: gb_un.*
- 15: gb_vl.*
- 16: em_ba.*
- 17: em_fun.*
- 18: em_hum.*
- 19: em_in.*
- 20: em_mu.*
- 21: em_om.*
- 22: em_or.*
- 23: em_ov.*
- 24: em_pat.*
- 25: em_ph.*
- 26: em_pl.*
- 27: em_ro.*
- 28: em_sts.*
- 29: em_un.*
- 30: em_vl.*
- 31: em_hlg_hum.*
- 32: em_hlg_inv.*
- 33: em_hlg_other.*
- 34: em_hlg_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query Match Length DB ID Description

1	2365	100.0	2365	6	AR157514	AR157514 Sequence
2	2365	100.0	2365	6	AR173203	AR173203 Sequence
3	2365	100.0	2365	6	HS4278110	HS4278110 Homo sapi
4	2188.4	92.5	2200	6	AX086506	AX086506 Sequence
5	2188.4	92.5	2200	6	HS4278110	HS4278110 Homo sapi
6	942.4	39.5	191137	9	AC010282	AC010282 Homo sapi
7	934.2	39.5	170569	2	AC108109	AC108109 Homo sapi
8	715.6	30.3	213447	2	AC083833	AC083833 Homo sapi
9	482.8	20.4	100061	2	AC016551	AC016551 Homo sapi
10	412.6	17.4	129261	2	AC068247	AC068247 Homo sapi
11	411	17.4	198052	9	AC019205	AC019205 Homo sapi
12	366.8	15.5	56864	9	AC010091	AC010091 Homo sapi
13	358	15.1	3270	3	AY051831	AY051831 Drosophila
14	307.8	13.0	323	6	AR173200	AR173200 Sequence
15	307.8	11.3	1780	8	VSPP680BP	VSPP680BP
16	267.8	11.3	1924	8	AB010259	AB010259 Homo sapi
17	266.4	11.3	2067	8	AF428313	AF428313 Arabidops
18	266.4	11.3	2165	8	AY050375	AY050375 Arabidops
19	266.4	11.3	2264	8	AY062351	AY062351 Arabidops
20	261.6	11.1	1814	8	AY062745	AY062745 Arabidops
21	257.2	10.8	2378	8	TOBRDB10	TOBRDB10
22	256.2	10.8	2415	8	AF462866	AF462866 Arabidops
23	235.2	9.9	2366	8	AY060589	AY060589 Arabidops
24	229	9.7	65957	2	AC014595	AC014595 Drosophila
25	229	9.7	170498	3	AC008094	AC008094 Drosophila
26	229	9.7	181950	3	AC008093	AC008093 Drosophila
27	229	9.7	227894	3	AE003677	AE003677 Drosophila
28	228.2	9.6	135853	9	AC010971	AC010971 Homo sapi
29	227.6	9.6	2318	10	MMHELI	MMHELI
30	227.6	9.6	2321	6	AX140497	AX140497 Sequence
31	226.4	9.6	203326	2	AL645849	AL645849 Mus muscu
32	224.6	9.5	2694	5	AF158370	AF158370 Gallus ga
33	222.2	9.4	2323	9	HSNP68M	HSNP68M
34	222.2	9.4	2366	9	BC016027	BC016027 Human mRN
35	222.2	9.4	2482	8	AY057700	AY057700 Homo sapi
36	221.8	9.3	82592	8	AC002337	AC002337 Arabidops
37	220.8	9.3	2739	3	HSPE8A	HSPE8A
38	220.6	9.2	2799	3	AF110009	AF110009 Molguila o
39	217.6	9.2	5374	3	AF110007	AF110007 Molguila o
40	217.6	9.0	5437	3	AF110008	AF110008 Molguila o
41	212.2	8.9	2733	8	SCPE6	X52649 S.cerevisia
42	211.4	8.9	3740	8	SCYNU112W	771388 S.cerevisia
43	211.4	8.9	37800	8	SCCYIV38K	Z69382 S.cerevisia
44	211.4	8.9	3740	9	AB018344	AB018344 Homo sapi
45	211	8.9	3740	9	AB018344	AB018344 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AR157514
DEFINITION Sequence 42 from patent US 6245525.
ACCESSION AR157514
VERSION AR157514.1 GI:16218459
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2365)
AUTHORS Martelange,V., De Smet,C. and Boon-Fallieur,T.
TITLE Tumor associated nucleic acids and uses therefor
JOURNAL Patent: US 6245525-A 42 12-JUN-2001;
FEATURES
source Location/Qualifiers
1..2365
BASE COUNT 775 a 420 c 592 g 578 t
ORIGIN

Query Match 100.0%; Score 2365; DB 6; Length 2365;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ttgttaccgagctcggaatccctagtaagcgccgctgctgctggaagaagtggcgggga 60
Db 1 ttgttaccgagctcggaatccctagtaagcgccgctgctgctggaagaagtggcgggga 60
OY 61 taagaagcgtgaggcgaggagcttaagcctcgctggggtcccttaagtaagcgtgctgagc 120
Db 61 taagaagcgtgaggcgaggagcttaagcctcgctggggtcccttaagtaagcgtgctgagc 120
OY 121 ttgcctgcaagcgttaagcgttaagcgttaagcgttaagcgttaagcgttaagcgttaagc 180
Db 121 ttgcctgcaagcgttaagcgttaagcgttaagcgttaagcgttaagcgttaagcgttaagc 180
OY 181 aagctggaagcgagccctctctggaacaaatgcccacacacggaagagctcccaagagctct 240
Db 181 aagctggaagcgagccctctctggaacaaatgcccacacacggaagagctcccaagagctct 240
OY 241 aagctggaagcgtgagctgagcgaagctgagcgaagctgagcgaagctgagcgaagctgagc 300
Db 241 aagctggaagcgtgagctgagcgaagctgagcgaagctgagcgaagctgagcgaagctgagc 300
OY 301 gcggaagagctggaatcgaaacaggtccctggaaggaataagtgctggaagagtgctgctgag 360
Db 301 gcggaagagctggaatcgaaacaggtccctggaaggaataagtgctggaagagtgctgctgag 360
OY 361 aagagcagctctagagcccgaggagcgctgagcgtgctgagcgaaggaagctgagcgtgctg 420
Db 361 aagagcagctctagagcccgaggagcgctgagcgtgctgagcgaaggaagctgagcgtgctg 420
OY 421 ttgcttgaagagcgaacttgctgagcgttaagcgtgctgagcgtgagcgtgagcgtgagcgt 480
Db 421 ttgcttgaagagcgaacttgctgagcgttaagcgtgctgagcgtgagcgtgagcgtgagcgt 480
OY 481 aatataaagatacaacaaaccccaatcccaatataatcaagaacaccccaatataatca 540
Db 481 aatataaagatacaacaaaccccaatcccaatataatcaagaacaccccaatataatca 540
OY 541 gtcaaaatcttgagcgaagcaagcaatgcaaacggaagcaaaagcagtgatgacaatctt 600
Db 541 gtcaaaatcttgagcgaagcaagcaatgcaaacggaagcaaaagcagtgatgacaatctt 600
OY 601 gctaaaaagctagaagaagaatcacatccagaaatgagcaatctgaactgcaactccact 660
Db 601 gctaaaaagctagaagaagaatcacatccagaaatgagcaatctgaactgcaactccact 660
OY 661 tctgttggaagaaatggaagcagacagatgacatgctgctgcaagagatcgccgcaatgata 720
Db 661 tctgttggaagaaatggaagcagacagatgacatgctgctgcaagagatcgccgcaatgata 720
OY 721 gattggatcaaatagagaagaaagttgaaatgcaaaataaaacaaagtgagcagattta 780
Db 721 gattggatcaaatagagaagaaagttgaaatgcaaaataaaacaaagtgagcagattta 780
OY 781 ccacaaatlaagaagaacatttataagaagatccactgccaagtgccatgcaaaagta 840
Db 781 ccacaaatlaagaagaacatttataagaagatccactgccaagtgccatgcaaaagta 840
OY 841 gaagcagatagctggaagaaagaagaatttaataataagctggaatgagcgttaagaagatg 900
Db 841 gaagcagatagctggaagaaagaagaatttaataataagctggaatgagcgttaagaagatg 900
OY 901 gaggaaacgacctatcccaatccactgcaaatctggaagaagccttcaatgctatcc 960
Db 901 gaggaaacgacctatcccaatccactgcaaatctggaagaagccttcaatgctatcc 960
OY 961 gaagctaaagaagaacaaagcgtttcaaaagcgaacacacacacacacacacacacac 1020
Db 961 gaagctaaagaagaacaaagcgtttcaaaagcgaacacacacacacacacacacacacac 1020
OY 1021 gcatggcccatgctgctgcaagaagatcatctatagagtgagccgagcagctggaacaga 1080
Db 1021 gcatggcccatgctgctgcaagaagatcatctatagagtgagccgagcagctggaacaga 1080
```

```
OY 1081 aagaatgctgcttaataagctggaattatcaatcgtgcttcaacccacacacacacacac 1140
Db 1081 aagaatgctgcttaataagctggaattatcaatcgtgcttcaacccacacacacacacac 1140
OY 1141 ggtcaaggaatlaagcccgcaatgctgcttaaccccaatccgaatgagcactcaaa 1200
Db 1141 ggtcaaggaatlaagcccgcaatgctgcttaaccccaatccgaatgagcactcaaa 1200
OY 1201 gtagaagaagaaatgcaaatctcataataagagcttggaatgctgctgctgctgctgct 1260
Db 1201 gtagaagaagaaatgcaaatctcataataagagcttggaatgctgctgctgctgctgct 1260
OY 1261 ggtggaatagatgagaaacaatagaagccttaaaaggtgtagatcatalcataatgca 1320
Db 1261 ggtggaatagatgagaaacaatagaagccttaaaaggtgtagatcatalcataatgca 1320
OY 1321 actccggaagatggaatgcatcgaatgagtaactcgctcaatctgaagaatataaac 1380
Db 1321 actccggaagatggaatgcatcgaatgagtaactcgctcaatctgaagaatataaac 1380
OY 1381 taactgcttgaatggaagcagaagaagatgctggaatggaatggaatggaatggaatgga 1440
Db 1381 taactgcttgaatggaagcagaagaagatgctggaatggaatggaatggaatggaatgga 1440
OY 1441 aagattgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1500
Db 1441 aagattgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1500
OY 1501 catcaagctcagcctcgacacacacacacacacacacacacacacacacacacacacac 1560
Db 1501 catcaagctcagcctcgacacacacacacacacacacacacacacacacacacacacac 1560
OY 1561 acattggaatcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1620
Db 1561 acattggaatcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1620
OY 1621 gagaagtggaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1680
Db 1621 gagaagtggaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1680
OY 1681 gctctgcttctcgaaagagctgctggaatcgaatcgaatcgaatcgaatcgaatcgaat 1740
Db 1681 gctctgcttctcgaaagagctgctggaatcgaatcgaatcgaatcgaatcgaatcgaat 1740
OY 1741 atataagtagagcttctgcaatggaatggaatggaatggaatggaatggaatggaatgga 1800
Db 1741 atataagtagagcttctgcaatggaatggaatggaatggaatggaatggaatggaatgga 1800
OY 1801 aactttaaagaagcagaagtggaatggaatggaatggaatggaatggaatggaatggaat 1860
Db 1801 aactttaaagaagcagaagtggaatggaatggaatggaatggaatggaatggaatggaat 1860
OY 1861 gattgctcagcgttcaacacacacacacacacacacacacacacacacacacacacac 1920
Db 1861 gattgctcagcgttcaacacacacacacacacacacacacacacacacacacacacac 1920
OY 1921 gtacacggaatggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1980
Db 1921 gtacacggaatggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1980
OY 1981 actagaagaatggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2040
Db 1981 actagaagaatggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2040
OY 2041 agtatccagaagagcctgctatcaatggaatggaatggaatggaatggaatggaatggaat 2100
Db 2041 agtatccagaagagcctgctatcaatggaatggaatggaatggaatggaatggaatggaat 2100
OY 2101 gaaatggaagaaagaatggaagaaatggaagaaatggaagaaatggaagaaatggaagaa 2160
Db 2101 gaaatggaagaaagaatggaagaaatggaagaaatggaagaaatggaagaaatggaagaa 2160
```

QY 2161 cgtactagtggtgagaaatcagaatcttctagaataatagtaagacagagatctg 2220
|||||
Db 2161 CTGACTAGTGGGTAGAGAAATTTTGTGAATATAGTAAAGACAGATATG 2220
QY 2221 gacatgttgagcagtaaaagacgagactgatttgactgattcttaataatagtgct 2280
|||||
Db 2221 GACATGTTGGCAGTATAGACAGCCGACTGATTGACTGATTCTTAATATATATGTGTT 2280
QY 2281 tgaataataagaatccagtgcttactacttcttactaataataagaatctaaact 2340
|||||
Db 2281 TGAATAATATAGATTCAGTGTCTTTATCTTTCTTAATATAATAGATTTAACTT 2340
QY 2341 aaaaaaaaaaaaaaaaaaaaaa 2365
|||||
Db 2341 AAAAAAAAAAAAAAAAAAAAAA 2365

RESULT 2
ARI73203 ARI73203 2365 bp DNA Linear PAT 17-DEC-2001
LOCUS Sequence 42 from patent US 6303756.
DEFINITION ARI73203
ACCESSION ARI73203
VERSION ARI73203.1 GI:17912694
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2365)
AUTHORS Martelange, Y., De Smet, C. and Boon-Falleur, T.
Tumor associated nucleic acids and uses therefor
JOURNAL Patent: US 6303756-A 42 16-OCT-2001;
FEATURES
source 1..2365
BASE COUNT 775 a 420 c 592 g 578 t
ORIGIN

Query Match 100.0%; Score 2365; DB 6; Length 2365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttagtaccagctcgcagatccctagtaacgcgcgcagctgcttgagaaagatgagcgga 60
|||||
Db 1 TTGGTACCGAGCTCGAATCCCTAGTAAAGCCGCCAGTGTCTGGAAGATGGCGGGA 60
QY 61 tagaagcgtggcgggggggtcagctcgtgctgggctccctaaagtcggtcgctggc 120
|||||
Db 61 TAGAGAGCTGGCGGGGGGCTAGCCTCGTGGGGCTCTTAACTAGCGGCTGCGTGGC 120
QY 121 ttcctggcagctagctcttaagcgtcagcgtcagctgctgagagcttgagcgcaacg 180
|||||
Db 121 TTCCCTGGCAGCTAGTCTTACGACGTCAAGGTCAAGTGTGCAAGCTGAGCGGCAAG 180
QY 181 aegtcgagcgcgccctctctggaacaatgtcccaacgagagagctcccaagcctct 240
|||||
Db 181 AGCTGGAGCGCCCTCTTCTTGAACAATGTCCCAACAGAGAGAGCTCCCAAGGCTCT 240
QY 241 acgtggtcgtctgctagtcgagcagctcgaacgtgtcccgagcgccagagagcg 300
|||||
Db 241 ACGTGGGTGTTGCTAGTCGCGGAGCTGACAGTGTCTCCGAGCGCCAGAGAGGCG 300
QY 301 gcgaggaagtcgaatcgaacaggtctcgaagataatgtctgcgcagagtgctgcctg 360
|||||
Db 301 GCGAGGAGATTGAATCGAACAAGTCTGAGAGGATATAGTGTGGAGAGGTGCTGG 360
QY 361 agagacacctcgaagcccccgagagcgctgagcgctgctcgaagaaatcgcgtgctg 420
|||||
Db 361 AGAGGACACTCTAGGCCCCCGAGAGCGCGTGGCGCTGCTCAAGAGAACTGCGGCTGT 420
QY 421 tctgcttgaaagacacatctgtctggcgcgtaactcgctggttggttcaaaaaataag 480
|||||
Db 421 TTTGCTTTTGAAGACCACTTTGTTGGCGCGTAAATCGTGTGGGTCAAAAAATAAAG 480

QY 481 aataacaaagtlacaaacaacacacacatccaaataatacaagaaacaaacgaatctla 540
|||||
Db 481 AATATACAAAGTACCAACCAACACCATCATCAATATATCAAGAACCAACGAATCAATTA 540
QY 541 gtcaaaatcttggcagcagagcgaaatgcaaaacgaagcaagcgatgtagaacaatttc 600
|||||
Db 541 GTCAAAATTTTTTGGCAGCAAGGC AATGCAAAACGAAGCAAAAGCAAGTGTATGACAAATTTT 600
QY 601 gtlaaaaagctagaagaataataacatccagaaatgcggaatlgatctgcatctcaacct 660
|||||
Db 601 GTTAAAGCTAGCAAGAAATTAACAATTCAGAAATGCGGAATGTATACGCAATCAACCT 660
QY 661 tctgttgaaagaatggaagcagatacaaatgtgttgcgagagatctgcgcattgata 720
|||||
Db 661 TCTGTGTAAGAAATGGAAGCAACATATACATATGTTGTTCAGAGAAATCGCGCATTTGTA 720
QY 721 gattggatcaaaatagaagaagaggttgaatctgcaaaacaaagctgagcagattta 780
|||||
Db 721 GATTGGATCAAAATTAGAGAGGAGGATTGAATGCGCAAAACAAAGTGGCGAGATTTA 780
QY 781 ccaccacatlaagaataaacttataaagatccacatgcgcacaaatgcatgtcaaaagta 840
|||||
Db 781 CCACCAATTAAGAAACCTTTATTAAGATCCACTGCCCAAGTGCATGTCAAAAGTA 840
QY 841 gaagcagatagctggaagaagaataatcttaataacgtggaatgacttgaagaatgag 900
|||||
Db 841 GAACGACATAGTTGAGAGAAAGAAATTTTAATTAACGTGGGATGATCTGAAGATGG 900
QY 901 gagaagacacctcccaatccctacgcgcacatttgaagaagcttcaatgataacct 960
|||||
Db 901 GAGAAAGACCTTATCCCAATCTTACCTGCAATTTATATACAGCCCTTCAATGTTATCT 960
QY 961 gaagttatcggaataacataaaagcaggttctcaaaagcgaacacatctacgtacacg 1020
|||||
Db 961 GAGGTTATGGAATAATTAATAAGCAGGTTTCAAAAGCAACACCTATTCAGTACAG 1020
QY 1021 gcatggcccatgtgtctgaagaataatcttctaaagatagccaagcttgaaacagga 1080
|||||
Db 1021 GCATGGCCCATGTGTTGCAAGCAATATATCTTATAGAAATACCCAGATGCAACAGGA 1080
QY 1081 aagacattgtcttaattaaatgcctgacatcttaacatctgtccctcaacccagccttaa 1140
|||||
Db 1081 AAGACATTGTTATTATTAAATGCTGATTTATTCATCTGCTTCACCCAGCCTTAA 1140
QY 1141 ggtcaagaagatagaccggcagtgatcttaactccatccatcggaatttaacattaa 1200
|||||
Db 1141 GGTCAAGGAATAGACCCGCAATGTAGTCTTAATCTCCACTCGGAAATTAACCTTAA 1200
QY 1201 gtagaagagaaatgtctgaataatcataaaagagctctcgagatgtctgtgataatg 1260
|||||
Db 1201 GTAGAAAGGAATGTTGCAAAATATTCATATTAAGGCGCTTCGAGGCTTTGTATATAGT 1260
QY 1261 ggttgaaatagagatgaataaataagaagctttaaagaaggttaatatctataatga 1320
|||||
Db 1261 GGTGAAATAGAGATGAACAAATAGAGAGGCTTAAAAAGGTGATATATCAATTAATGCA 1320
QY 1321 actcccggaagatlgatgaltctgcaaaatgaaatctgctcaatctgaaataataac 1380
|||||
Db 1321 ACTCCGGAAGATGAAATGATCTGCAAAATGATGAACTTCTCAATCTGAAGAAATTAAC 1380
QY 1381 taactgtctttaaagtagaagcagaagatgttgaacatggatcttaaacccaagtaatg 1440
|||||
Db 1381 TACTGTGTTTAAATGAAGAGACAAAGATGTTGACATGGGAATTTGAACCCAGATTAAG 1440
QY 1441 aagatctgtctgaatgtgcgcacaaatagagcagaatctgaacagtgctcaatgagcct 1500
|||||
Db 1441 AAGATTTTGTAGATGTGCGCCAGATAGCGCAGACATTAATGACCGAGTCTACATGGCT 1500
QY 1501 catcagttcaatgcctcgcaaatcttaattgaanaacaaatgattctcatgttgt 1560
|||||
Db 1501 CATTCAGTTATGCGCTCGCAATCTTATTTGAAGAACCAATGATGTCTATGTGTGT 1560

QY 1561 acattggaatcagtcgtgaagtcgaagcaaaatataatctgaaccacgagga 1620
DB 1561 ACATTGGAATCAGTCGTGAAGTCGAAGCAAAATATAATTTGAACACCGAGGA 1620
QY 1621 gagaaatggaatcagtcgaatcttctacagatctgcatcacaagaatcatt 1660
DB 1621 GAGAAATGGAATCAGTCGAATCTTCTACAGATCTGCTCATCCAGCAAGTCATT 1660
QY 1681 gcttcgtcttcgaaagcgtgtgagagacatcatcaagtcgaatctgaat 1740
DB 1681 GCTTCGTCTCTCGAAAGCGTGTGAGATCATCTTATCAAGTCACTTACTTGAAAT 1740
QY 1741 atataagtagagtccttcgacagagatagaagaagagagagagagagagag 1800
DB 1741 ATATCAGTAGAGTCCTTCGATGAGATAGAGAACAGAGATCGGGAAGCAATTAGAG 1800
QY 1801 aactttaaacaagcaaatgagaaataactatcaactgacatccctcagagact 1860
DB 1801 AACTTTAAACAGCAAAAGTGAAGTAATGCAATTCGCACTGATCTAGAGACTT 1860
QY 1861 gatgtccatgacgttcaacatctctataatctgaatctcacaagaaatcga 1920
DB 1861 GATGTCCATGACGTTACACATGCTTAATTTTACCTTCCACGGAATTTGAGAAATAC 1920
QY 1921 gtacacccgaatagagcgacaggaagagagagagagagagagagagagag 1980
DB 1921 GTACACCGAATAGCGCGCACGGGAAGACGAGAGAGAGCTGTCTTCAATTCACCTTTG 1980
QY 1981 actagaatagatgagagaggtgctccctgaatgataatctcagaagaagca 2040
DB 1981 ACTAGAAATGATGAGAGGCTGCTCTGAATTTGATTAATTTCTGGAAGACCAATATAG 2040
QY 2041 agtattccagagagagctgtatcaatgagctgagagagctgagagacacac 2100
DB 2041 AGTATTCAGAGAGAGCTGTGATCAATGGCTGAGAGGTTTGAAGCAGATCAGAGAAAGG 2100
QY 2101 gaaatggaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2160
DB 2101 GAAATGGAAGAAGAAGAAGAAGACCTCAAGAGAGGCCCAAGAGTTTCAATTAAGTCTT 2160
QY 2161 ctgtactagtcgagtagagaatcagaatcttcttgaagaatagtagaaga 2220
DB 2161 CTGTACTAGTCGAGGTAGAGAAATTTTGAAGATTAATGTAAGACAGAACTATTTG 2220
QY 2221 gacatgttcgagtagagaagaagaagaagaagaagaagaagaagaagaaga 2280
DB 2221 GACATGTTGCGAGTATGAGAGACCGGACGATTTGACTGATTTTAAATTAATGCTTT 2280
QY 2281 tgaataatagaaatcagtcgttcttacttcttcttcttcttcttcttctt 2340
DB 2281 TGAATAATAGAAATCCAGTGTCTTACTTCTTAAATAAATAGAACTATTTAACTT 2340
QY 2341 aaaaaaaaaaaaaaaaaaaaaa 2365
DB 2341 AAAAAAAAAAAAAAAAAAAAAA 2365

RESULT 3
HSA278110 2365 bp mRNA linear PRI 26-SEP-2000
LOCUS HSA278110
DEFINITION Homo sapiens mRNA for DEAD-box protein (HAGE gene).
ACCESSION AJ278110
VERSION AJ278110.1 GI:8216984
KEYWORDS dead-box protein; HAGE gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2365)
AUTHORS Martelange V., De Smet C., De Plaen E., Lurquin C. and Boon T.
TITLE Identification of a human sarcoma of two new genes with
tumor-specific expression
JOURNAL Cancer Res. 60 (14), 3848-3855 (2000)

★

MEDLINE 20374312
REFERENCE 2 (bases 1 to 2365)
AUTHORS Martelange V.M.F.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2000) Martelange V.M.F., Ludwig Institute for
Cancer Research, Université Catholique de Louvain, Av Hippocrate,
74, 1200 Bruxelles, BELGIUM
FEATURES
source
1..2365
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="q12-q13"
/cell_line="LB23-SAR"
/cell_type="rhabdomyosarcoma"
/note="marker GAT11808-D6S284"
208..2154
/gene="HAGE"
208..2154
/gene="HAGE"
208..2154
/function="putative RNA helicase ATP dependent"
/product="DEAD-box protein"
/protein_id="CAB92442.1"
/db_xref="GI:8216985"
/translation="MSHHGCAKASTWVVASRSTYSRAPERRAEELNRTGEGTS
VGRGRRKRTSRPEAVAGHEELPLCFALSKHFGAVTGRGSKIKNTOSTWTTIO
IIIDQPSLVKIFGSKAMOTKAAVIDNHFVKLEENNSGCIIDAFQPSKDSSTND
NNVAVGDRPLIDMDQIREGLKMOKTWADLPPIKKNFKESTASANSKYBADSMRK
ENFNITWDDLDKDEKRPINPTCFEDDAFCQPEWMENTKAGFOKPPIOSQAMPY
LOGIDILGVNQTGKTLCLYLMPEFRIHLVLOPSLKGORNPGLVLPTELALQVIG
ECKSYSGKLSVYCGGNGRDEQIEELKGVDDIIATPERLNDLQMSFNFLNITY
LVDEADKMLDMGFEPOIMKILIDVRPDRQVWMSATWPSVHRLAOSYLEPNTIVY
GTDILVAVSSYKONILTYPEEFKSHQPTLOSMSYSDKVIYVSRKAYADHSDDLI
LGNTSVSLSDGREDQRDRKALENFKKRLILADLDSRLDVIHDVTHYINDFPR
NIEIVRIKRTGRAGRTVSTITLIRDMNVAVSELINILIRANOSIPBELVSAENR
EAHORKEKMRKMERPOGPKKFKH"
BASE COUNT 775 a 420 c 592 g 578 t
ORIGIN
Query Match 100.0%; Score 2365; DB 9; Length 2365.
Best Local Similarity 100.0%; Pred No. 0;
Matches 2365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tgggtacgaagctcgatccctagtagtaacgagcgagtgctgagaaagatggcgagga 60
DB 1 TTGCTACCGAGCTCGATCCCTAGTAACGGCGCCAGTGTCTGTGAAAGAGTGGCGGGA 60
QY 61 tagaagagctggcgagggagagctgctgctgagagctcccttaagtagcgagctgagc 120
DB 61 TAGAGAGCTTGGCGGAGGAGCTCTGCTGCGGCTCTTTAAGTAGCGGCTGCGTGGC 120
QY 121 ttccttcgacgctagcttcaagcgtcagctgagtgagtgagcagcctgagcagcag 180
DB 121 TTCCTTCGACGCTAGCTTACGACGTCAGCTGAGTGTGTCAGAGACTGAGCGCAACG 180
QY 181 acgttcgagcgccctctctggaacaatgctccacaacgagagagcttcccaagcctct 240
DB 181 ACCTGCGAGCGCCCTCTTCTTGGAACAATGTCCACACGAGAGAGCTCCCAAGGCTCT 240
QY 241 acgttgagctgtctgtagtcgagcagctcgaagctgctccagcagcagagagagcgc 300
DB 241 ACCTGAGCTGTCTGCTAGTGTGCGGAGCTCGACAGTGTCCGAGCGCCAGAGAGAGCGCG 300
QY 301 ggcgaagagtgaaatcgaacaggttcctggagagataatgtcgcgcagagatgctgctg 360
DB 301 GCGGAGAGTTGAATCGAACAGGCTCTGAGGAGTATAGTGTGCGCAGAGTGTGCTGCG 360
QY 361 agagagacacctgaagcccccagagagcgctgagcagctgtagcagaagaatgcagctg 420
DB 361 AGAGAGACCTTAGAGCCCGGAGAGCGCTGCTGCTGCTCAGAGAACTGCGCGTGTGT 420

TITLE Human dna sequences
JOURNAL Patent: WO 0112653-A 458 22-FEB-2001:
German Human Genome Project (DE)
FEATURES
Source Location/Qualifiers
1. .2200
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 756 a 377 c 523 g 544 t
ORIGIN

Query Match 92.5%; Score 2188.4; DB 6; Length 2200;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 176 caacgaacgtcgaacgcgcgcctcttggaacacatgcccacccggaaggagctcccaag 235
Db 1 CACGACCTCGAGACGCCCTTTCTTGACACATGTCACACCGAGAGAGCTCCCAAG 60
Qy 236 cctctaegtggtctgtcgtcagtcggaagctcgaacgtgtcccgagcgcaagaaga 295
Db 61 CCTCTACGTGGTCTGTTGCTAGTCGGCGAAGCTCGACAGTGTCCGACGCCAGAGAGA 120
Qy 256 ggcgcgcggaagagttgaaatcgaaacagg tccctgaaggatatagtgtcggcaagg tgc 355
Db 121 GGCCGCGGAGAGATGTAATCGAACAGTCTTGAGGGAATATAGTGTGGCAGAGAGTGTGTC 180
Qy 356 gctggaagaggaacccctcagcccccggaagcgctgagcgtgtcagcgaacgcgcgc 415
Db 181 GCTGAGAGGACCTCTAGGCCCGGAGGCCGTGGCCGCTGTGTCAGAGAACTGCCGC 240
Qy 416 tgtgttctgcttgaagagcactctgtctgagcggaatcggctggtggtggtcaaaa 475
Db 241 TGTGTTTGTCTTGAAGAGCCACTTGTGTGGCGGGAATCGTCTGTGTGGTCAAAA 300
Qy 476 taaagaatatcaaaagtacaaacaccacaatccaaataatacaagaacaacgaat 535
Db 301 TAAAGAAATATACAAAGTACACAAACACCAATCCAAATATATCAAGAACACCGAAT 360
Qy 536 catlagtcaaaatctcttgacgaagaagcaatgcgaacgaagaagaagaagtgatagaca 595
Db 361 CATTAGTCAAAAATTTTGGCAGCAGGCAATGCAGAAAGCAAAAGCAGATGAGACA 420
Qy 556 attctgttaaaagccaagaagaattacaatctcaaga tgcggaattgtactgtcattcc 655
Db 421 ATTTTGTAAAAAGCTAGAAAGAAATTTCAATTCAGATCGGAATTTGATCTGATTC 480
Qy 656 aaacctctcttgaaaaaga tgaagcacagatacaaa tgtgtgtcagagaaga tgcgcat 715
Db 481 AACCTCTCTGTGGAAGAAAGATGGAACACAGATATACATTTGTTGAGAGAGATCGGCAT 540
Qy 716 tgaatagcttgggatacaaatlagagaagaagttgaaatgycaaaaaacaagtyggcag 775
Db 541 TGATAGATTGGATCAAAATTAGAGAGAAAGTTGAATTTGCAAAAACAAAGTGGCAG 600
Qy 776 atttacaccaatlaagaaanaactttaaagagtcacttgcacaaagtgcacgtcaaa 835
Db 601 ATTATCACCAATTAAGAAAACTTTTATAAGAGTCCACTGCCACAAAGTGCATGTCAA 660
Qy 836 aagtaagaacagatagctggaggaagaanaatttaatacaacgttggagacttgaaag 895
Db 661 AAGTAGAACGAGATAGTTGGAGAAAGAAATTTTAATATTAACGTGGATGATCTGAAG 720
Qy 896 atggggaagaagaacatacccaactcctaccgtgcacatttgaagccttcaagtc 955
Db 721 ATGGGGAAGAACACTATCCCAATCTCTACCTGCAATTTGATGACGCTTTCAATGTT 780
Qy 956 atcctgaagttatcgaaaacaataaaaagcaggtttcaaaagccaacacta ttcagt 1015
Db 781 ATCTGAGGTTATGGAAGAAACATTAAGAAAGGAGGTTTCAAAAGCCAACACCTATTCAGT 840
Qy 1016 caaaggaatggccctctgtgtgcaaggaatagatctttaagagtagccaaagctgaa 1075

Db 841 CACAGCAGTGCCCAATGTGTGCAAGCAATAGATCTTATAGGAGTACCCAGACTGGAA 900
Qy 1076 caggaaagaacatgtgtatcttaatgcctggaatctcaatcgttccctcaaacagcc 1135
Db 901 CAGGAAGACATTTGTTATTAAAGCCGTGATTTATCATCTGTGCTTCAACAGCC 960
Qy 1136 ttaaaagtcaaaaggaatagaccgcgca tfttaatttaacttccactccgggaattgac 1195
Db 961 TTAAGGTCAAAGAAATAGACCCGCAATGTTAATCTTAACTCCACCTCGGAATTAGCA 1020
Qy 1196 tcaagtagaaggaagatgttcgaatctcataataaagagctcgaagtgtgtgcat 1255
Db 1021 TTCAAAGTAAGCAATAGTTGCAAAATTTATATTAAGGCTTGAGGTGTTGTGTAT 1080
Qy 1256 atgtgtgtgaaatagaatagaacaaatagaagaagcttaaaaagtgtagatactaa 1315
Db 1081 ATGGGTGGAATATAGATAGAACAAATAGAAAGCTTAAAAAAGCTGTATATATCTAA 1140
Qy 1316 tgcgaactcccggaagatctgaatctgcaaaatgaatctgtccaatctgaagata 1375
Db 1141 TTGCAACTCCCGAAGATTGAATGATCTGCAAAATGAGTAACTGTCAATCTGAAGATA 1200
Qy 1376 taactacttgttttaagaagaacagaagaagtgttgacctgggaattgaaacccaga 1435
Db 1201 TTAACACTTGTGTTTATAGTAAAGCAGACAAAGATGTTGACATYGGGATTTGAACCCAGA 1260
Qy 1436 taaatgaagatcttctgaatctgcgccagaatagagaacagttcgaacagtgatcata 1495
Db 1261 TAATGAAAGATTTTGTATGATGTGCCGCCAGATAGGAGACAGATTATGACAGAGTACAT 1320
Qy 1496 ggcctcactcaatctatccgtctgcgaacatctctatttgaagaacaaatgattgtctatg 1555
Db 1321 GGCCCTATTCAGTTATATGCTCTGCAAAAGCTGTTGCGGATCACCTTATCAAGTACCTTA 1380
Qy 1556 tctgtacatctgaatcagttctgtgaagttcagttgaagcaaaataataatgttaacaacg 1615
Db 1381 TTGTGACATTTGGATCTAGTACTGCTAGTACAGTACCAAAATATATTTGTAACACCG 1440
Qy 1616 aggaagaagaatggaagtacacatgcaaatcttctcaagatgtctcaacagaagaag 1675
Db 1441 AGCAAGAGAAATGGAGTACATGCAAACTTTTCTACAGGATGTGCTACCAAGCAAGAAAG 1500
Qy 1676 tcatgtctctgtctctcgaaaagctgttgcgatacctta tcaagtgacctaaactgtg 1735
Db 1501 TCATTTCTCTCTCTCTCTGCAAAAGCTGTTGCGGATCACCTTATCAAGTACCTTA 1560
Qy 1736 gaaatatcaatgaagtcctctgcattgagatagaagaacaagaagatccggaagaagcat 1795
Db 1561 GAAATATATCAGTAGAGTCTCTGCAATGAGATAGAACAGAGATCGGAGAAAGCAT 1620
Qy 1796 tagagaacttcaaaacaggaagaagtgaagaatactaatgtcaactgtactagacctcagaag 1855
Db 1621 TAGAGAACTTTAAACAGGCAAGAGTGAATACTAATTTGCAAGTGAATCTAGCTTAGAG 1680
Qy 1856 gacttgatgtccatgaagttacacatgctataatcttggacttccacggaataatctgaag 1915
Db 1681 GACTTGATGTCCATGACGTTTACACATGCTCTAATTTTGTGACTTCCAGGAAATTTGAAG 1740
Qy 1916 aatacgtiacacggaatagggcgca cgggaagaacaggaaggaagctgtgtctcatactaca 1975
Db 1741 AATAGCTACACGGAATAGGGCGCACGGGAAGACAGGAGACATGCTTTTCATTACAA 1800
Qy 1976 ccttgactcagaagaatgacttggagagtgctcctcgaattgacttaataatctcgaagaagcaa 2035
Db 1801 CTTTGACTTGAAGAAATGATTGAGGGTTGCTGTGAATTTATTAATTTCTGAAAGAGCAA 1860
Qy 2036 atcaagatctccagaagaagctgtcta tcaa tggcctgaagaagtttgaagcaatccaaacgga 2095
Db 1861 ATCAGAGTATTCAGAGAGACTTGTATCAATGCTGAGAGGTTTGAGGCAATCAACAGGA 1920
Qy 2096 aagaggaata tggaaagaanaa tggaaagactcaaggaagggccaaagattcattaat 2155
Db 1921 AAAGGAAATGGAAGAAATGGAAGACCTCAAGAGAGGCCCAAGAAATTTCAATTAAT 1980

Qy 22156 gctcttcctgacacagcaggggctagaagaaatcttcttaaaaaataagaaagacagag 22155

Db 1981 gctttcttgactactgagggcgtgacgaattcmaagtttttttgaagattagtaacacagag 2040

Qy 2216 tactgacacatgctgcacgactagaagacccgacactgaatttactgaattcttaaaataa 2275

Db 2041 tatggacatggtggcagatgtaagagaccggacgtgattgactgattttaaataata 2100

Qy 2276 gctgttggaaataatagaatccagtgcttctatccttcttbaataaaatagaagtta 2335

Db 2101 gctttttgaaaataatgaaatccagctgattttatcttcttttaataaaaattgaaagatttta 2160

Qy 2336 aacttaaaaaaaaaaaaaaaaaaaaaa 2365

Db 2161 aacttcaaaaaaaaaaaaaaaaaaaaaa 2190

RESULT 5	2200 bp	MRNA	linear	PRI 10-MAR-2001
HSR801719				
Human sapiens mRNA, cDNA DKFZP434H2114 (from clone DKFZP434H2114); complete cds.				
AL136751				
AL136751.1	GI:12053018			
ORGANISM	human.			
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.			
AUTHORS	1 (bases 1 to 2200) Wiemann, S., Well, B., Wellenreuther, R., Gassenhuber, J., Glassl, S., Ansorge, W., Boecher, M., Blocher, H., Bauersachs, S., Blum, H., Lauber, J., Duesterhoeft, A., Beyer, K., Koehrer, K., Strack, N., Mewes, H.-W., Ottensmeyer, B., Obermayer, B., Tampe, J., Heubner, D., Wambutt, R., Korn, B., Klein, M. and Poustka, A.			
TITLE	Toward a Catalog of Human Genes and Proteins: Sequencing and Analysis of 500 Novel Complete Protein Coding Human cDNAs			
JOURNAL	Genome Res. 11 (3), 422-435 (2001)			
PUBMED	11840166			
REFERENCE	2 (bases 1 to 2200) Koehrer, K., Beyer, A., Mewes, H.-W., Gassenhuber, J. and Wiemann, S.			
AUTHORS	Direct Submission			
TITLE	Submitted (18-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY			
JOURNAL	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@kitz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Charité, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.			
COMMENT	This clone is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://www.mips.biochem.mpg.de/proj/cDNA/ . Location/Qualifiers			
FEATURES	1. 2200			
source	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="DKFZP434H2114" /tissue_type="testis" /clone_lib="434 (synonym: htss3). Vector popttl; host DH10B; sites NotI + SalI" /dev_stage="adult" 33. 1979 /gene="DKFZP434H2114" 33. 1979 /gene="DKFZP434H2114" /note="strong similarity to RNA helicases" /codon_start=1 /product="hypothetical protein" /protein_id="CAB86685.1" /db_xref="GI:12053019"			

[illegible]

QY 956 atccgaggtatctgaaacacatctaaaggcaggtcttccaaagccacacatctacat 1015
 |||||
 Db 781 ATCTGAGGTTATGGAACATTTAAAGGCGAGTTTCCAAAGCCACACCTATTCCAGT 840
 QY 1016 caacgagcatgcccacatctgctgctcaaggaatagatctctatagagtaagcccaagctgaa 1075
 |||||
 Db 841 CACAGGCAATGCCCCATTGCTGTGCAAGGAATAGATCTTATAGAGGTAGCCCCACACGGA 900
 QY 1076 caggaagaacatctgttacttaattacgtccgtgagttatccatctgctcctcaagccagcc 1135
 |||||
 Db 901 CAGGAAGACANTGCTGTTATTATTATGCTGATTTATTCATCTGCTGCTTCAACCCAGCC 960
 QY 1136 taaaggtccaaaggaatagaccgcgagatcttagtctcctaactcccaactcgggaattagca 1195
 |||||
 Db 961 TTAAAGGTCAAAGGAATAGACCCGCGATGTTAGTCTTAACCTCCACCTCGGGAATTAGCAC 1020
 QY 1196 tccaagtagaaggaatgctctgcaaatctcaataatcaagaaggtccggagctgtctgtat 1255
 |||||
 Db 1021 TTCAAGTAGAAGGAGATGTTGCAATATTCTATTAAGGGCTTCGAGAGTGTGCTGAT 1080
 QY 1256 atggtggtggaatagagatgaaacaaatagaagagcttaaaaaagtgtagatacata 1315
 |||||
 Db 1081 ATGGTGTGGAATAGAGATGAAACAATAGAAAGCTTAAAGGCTGTAGATATCATTA 1140
 QY 1316 tctgaactcccggaagaatgaaatgacatcgaatgaactctgcaactctgaagaata 1375
 |||||
 Db 1141 TTGCAACTCCCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 QY 1376 taactactgtgtttatagatgaaagcaagaatgctgaaatgaggtctgaagcccaaga 1435
 |||||
 Db 1201 TAACTACTGTGTTTATGATGAAAGACAGACAGATGTTGCAATGGAATTGAAACCCCA 1260
 QY 1436 taatgaagaatctgttagatgctgcccagatagacacaggtatgaccagctgctacat 1495
 |||||
 Db 1261 TAAATGAATTTTGTAGATGTCGCCCGCAGATAGCAGACAGTATGACAGTGTACAT 1320
 QY 1496 ggcctcattcagttcatcgcgtcgcgaacatctctatctgaagaagcccaatgctgctatg 1555
 |||||
 Db 1321 GGCCTCATTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 QY 1556 tctgtacatctgagatcagttctgctgaagttcagtgaaagcaaatatgtaaccacg 1615
 |||||
 Db 1381 TTGGTACATTTGGATCTGTTGCTGTTAGTTGAGTTCAGTCAGTCAAAATATGTTAACCCG 1440
 QY 1616 aggaagaagaatgagatgacacatgcaaaccttctcaagaagatgctgtaaccagaagaag 1675
 |||||
 Db 1441 AGGAAGAGAAATGAGTCACATGCAAACTTTCTACAGAGATGTCATCCACAGCAAAAG 1500
 QY 1676 tcaatgctctcgtctctcgaaagcgtctgcgatcactatcaagtgaactaactgtg 1735
 |||||
 Db 1501 TCATTGCTCTTGTGTTTCGAAAGCTGTTGCGGATCATCTTATCAAGTGACCTTATACTTG 1560
 QY 1736 gaaatatactagtagtctctcgaatgagaatagagaagcaaggaagcaaggaagcaagc 1795
 |||||
 Db 1561 GAAATATATCAGTACAGTCTCTCTGATGAGATAGAGAACAGAGATCGGGAAGAAAGCT 1620
 QY 1796 tagagaactttaaaacagcgaaagtgaataactaattgcaactgtaactagcctcagag 1855
 |||||
 Db 1621 TAGAGAACTTTAAACAGCGAAAGTGAATCTTAAATGCAACTGATCTAGCCTCTAGAG 1680
 QY 1856 gactgagctcgaatgacatgctacacatgctataatcttgaaattccacgaatctgagag 1915
 |||||
 Db 1681 GACTTGATGTCATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 QY 1916 aatacgtatacccgaaatagggcgcaaggaagaagcaaggaagcaaggaagcaaggaag 1975
 |||||
 Db 1741 AATACGTATACCGAATAGGCGGACGGAAGAGCAGGAGAGACTGCGTGTTCATTAA 1800
 QY 1976 ctctgactagaatgactgaggtgtgctcctgaaatgatatctcgtgaagaagca 2035
 |||||
 Db 1801 CTTTGACTAGAAATGATGAGGCTTGCTGCTGATTTGATTTATTTCTTGGAAGAGCAA 1860

QY 2036 atcagaatctccagaagagctctgtatcatatgctcgtgagaggtttgaagccacatcaagga 2095
 |||||
 Db 1861 ATCAGAGTATTTCCAGAGAGCTTGTATCAATGAGCTGAGAGCTTGAAGCAGCATCAACGGA 1920
 QY 2096 aaaaggaaatctgaaagaaatctgaaagaccctcaaggaagcccaagaagttcctaact 2155
 |||||
 Db 1921 AAAGGAATATGAAAGAAATATGAAAGACTCAAGGAGAGCCCAAGAAAGTTTCATTAA 1980
 QY 2156 gctctgctactgtagggtagaagatctcaagaattcttagaataatgtaagaagaag 2215
 |||||
 Db 1991 GCTCTTCTGTCATGTCGTGAGTACAGATTTTGAAGATATGATAGACAGAGAG 2040
 QY 2216 taatgacatctgtgcagatataagagaccggaactgatactgactgattcttaaaataa 2275
 |||||
 Db 2041 TATTGACATGTTGGCAGTATGAAGACACCGAGCTGATTGACTGTTCTTAAATTA 2100
 QY 2276 gtgttgaaataatagaaaccagctcttataactctcttaataaataagaagattat 2335
 |||||
 Db 2101 GTGTTTAAATATATGAAATCCAGTGTGTTTATCTTTCTTAAATATGAAATATTA 2160
 QY 2336 aacttaaaaaaaaaaaaaaaaaaaaaa 2365
 |||||
 Db 2161 AACTTGAAAAAATAAAAAAAAAAAAAA 2190

RESULT 6
 AC010282
 LOCUS AC010282 191132 bp DNA linear PRI 01-SEP-2000
 DEFINITION Homo sapiens chromosome 5 clone CTC-55119, complete sequence.
 AC010282
 VERSION AC010282.5 GI:9958012
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 191132)
 TITLE DOE Joint Genome Institute and Stanford Human Genome Center.
 JOURNAL
 Direct Submission
 2 (bases 1 to 191132)
 TITLE DOE Joint Genome Institute.
 JOURNAL
 Direct Submission
 Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 3 (bases 1 to 191132)
 TITLE DOE Joint Genome Institute and Stanford Human Genome Center.
 JOURNAL
 Direct Submission
 Submitted (01-SEP-2000) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 On Sep 1, 2000 this sequence version replaced gi:9256180.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 COMMENT
 finishing Completed at Stanford Human Genome Center
 www.shgc.stanford.edu
 Quality: Pipap Quality >=40 99.2% of Sequence;
 Estimated Total Number of Errors is 1.5.
 STS Content:
 SHGC-100210 G55156
 WI-3394 G02870
 WI-11812 G24004
 SHGC-105345 G58760.

FEATURES
 source
 1. 191132
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="CTC-55119"
 BASE COUNT 61612 a 37225 c 35694 g 56601 t
 ORIGIN

Query Match

39.6%; Score 937.4; DB 9; Length 191132;

URES
source

Query Match	30.3%	Score 715.6;	DB 2;	Length 213447;
Best Local Similarity	67.5%;	Pred. No. 6.7e-143;		
Matches 937; Conservative	0;	Mismatches 449;	Indels 2;	Gaps 1;

[illegible]

QY	1290	gcttaaaaaagctgtaatatcatctaacctccggaagattgaatgctctgcaat	1349
Db	90283	CCTTAGAAGAGAGAAATATCATATTTGGCAACTCCCAAGAGCTGATGATCTCAAT	90224
QY	1350	gatttaactctgcataactctgagaataataacctacttgcttgatgagaaagcaagaat	1409
Db	90223	GAAATACFTTGTGCAGCTGAGAGTATACCTACTGCTGTTTATGAATAAGCAGACATAT	90164
QY	1410	gttgagaactgagattttaaaccagaataataagaatttctgatactgctgcagataag	1469
Db	90163	GCTGAGACATGGCGGTGTAACCCAGATATATGAAAGATTTGTATNNNNNNNNNNNNNN	90104
QY	1470	gcagacagctatgacccagctgctacatgagccctcaatcagttcaatgcctctgcacaactcta	1529
Db	90103	NN	90044
QY	1530	tttgaagagaccaatgactgcta	1553
Db	90043	NNNNNNNNNNNNNNNNNNNNNNNA	90020

RESULT	10
AC068247	
LOCUS	
DEFINITION	AC068247 129261 bp DNA linear HTG 09-JUN-2006
ACCESSION	AC068247
VERSION	AC068247.2 GI:8389573
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 129261)
TITLE	Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL	Homo sapiens, clone RP11-592J5
REFERENCE	unpublished
AUTHORS	2 (bases 1 to 129261)
	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

TITLE	Direct Submission
JOURNAL	Submitted (30-Apr-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Jun 9, 2000 this sequence version replaced g1.7671314.

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information -----
 Center project name: L7465


```
----- Summary Statistics -----
Center clone name: 592.J-5
Sequencing vector: M13; M77815: 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 111924 bases at least Q40
Consensus quality: 120625 bases at least Q30
Consensus quality: 124069 bases at least Q20
Insert size: 182000; agarose-fp
Quality coverage: 2.7 in Q20 bases; sum-of-coverage
Quality coverage: 3.9 in Q20 bases; sum-of-coverage

* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1020 1119: gap of 1019 bp in length
1120 2233: contig of 1114 bp in length
2234 2333: gap of 100 bp
2334 3388: contig of 1055 bp in length
3389 3488: gap of 100 bp
3489 4745: contig of 1257 bp in length
4746 4845: gap of 100 bp
4846 7126: contig of 2281 bp in length
7127 7226: gap of 100 bp
7227 9254: contig of 2028 bp in length
9255 9354: gap of 100 bp
9355 10467: contig of 1113 bp in length
10468 10567: gap of 100 bp
10568 12043: contig of 1476 bp in length
12044 12143: gap of 100 bp
12144 13675: contig of 1532 bp in length
13676 13775: gap of 100 bp
13776 15207: contig of 1432 bp in length
15208 15307: gap of 100 bp
15308 17676: contig of 2369 bp in length
17677 17776: gap of 100 bp
17777 20366: contig of 2590 bp in length
20367 20466: gap of 100 bp
20467 23113: contig of 2647 bp in length
23114 23213: gap of 100 bp
23214 25686: contig of 2473 bp in length
25687 25786: gap of 100 bp
25787 28215: contig of 2429 bp in length
28216 28315: gap of 100 bp
28316 30549: contig of 2234 bp in length
30550 30649: gap of 100 bp
30650 33299: contig of 2650 bp in length
33300 33399: gap of 100 bp
33400 36023: contig of 2624 bp in length
36024 36123: gap of 100 bp
36124 37903: contig of 1780 bp in length
37904 38003: gap of 100 bp
38004 41092: contig of 3089 bp in length
41093 41192: gap of 100 bp
41193 45421: contig of 4229 bp in length
45422 45521: gap of 100 bp
45522 48381: contig of 2860 bp in length
48382 48481: gap of 100 bp
48482 54521: contig of 5840 bp in length
54522 54421: gap of 100 bp
54422 59074: contig of 4653 bp in length
59075 59174: gap of 100 bp
59175 62454: contig of 3280 bp in length
62455 62554: gap of 100 bp
62555 66890: contig of 4336 bp in length
66891 66990: gap of 100 bp
```

```
FEATURES
source
* 66991 70314: contig of 3324 bp in length
* 70315 70414: gap of 100 bp
* 70415 77680: contig of 7266 bp in length
* 77681 77780: gap of 100 bp
* 77781 82561: contig of 4781 bp in length
* 82562 82661: gap of 100 bp
* 82662 91256: contig of 8595 bp in length
* 91257 91356: gap of 100 bp
* 91357 99646: contig of 8290 bp in length
* 99647 99746: gap of 100 bp
* 99747 115414: contig of 15668 bp in length
* 115415 115514: gap of 100 bp
* 115515 129261: contig of 13747 bp in length.
Location/Qualifiers
1..129261
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-592J5"
/clone_id="RP11 Human Male BAC"
1..1019
/note="assembly-fragment"
1120..2233
/note="assembly-fragment"
2234..3388
/note="assembly-fragment"
3389..4745
/note="assembly-fragment"
4746..7126
/note="assembly-fragment"
7127..9254
/note="assembly-fragment"
9255..10467
/note="assembly-fragment"
10468..12043
/note="assembly-fragment"
12044..13675
/note="assembly-fragment"
13676..15207
/note="assembly-fragment"
15208..17676
/note="assembly-fragment"
17677..20366
/note="assembly-fragment"
20367..20466
/note="assembly-fragment"
20467..23113
/note="assembly-fragment"
23114..25686
/note="assembly-fragment"
25687..28215
/note="assembly-fragment"
28216..30549
/note="assembly-fragment"
30550..33299
/note="assembly-fragment"
33300..33399
/note="assembly-fragment"
33400..36023
/note="assembly-fragment"
36024..37903
vector_side:right
36124..37903
/note="assembly-fragment"
38004..41092
/note="assembly-fragment"
41093..45421
/note="assembly-fragment"
45422..48381
/note="assembly-fragment"
48382..54321
/note="assembly-fragment"
54422..59074
/note="assembly-fragment"
59075..62454
/note="assembly-fragment"
62455..66890
/note="assembly-fragment"
66891..66990
/note="assembly-fragment"
```

```

misc_feature      6691..70314
                  /note="assembly_fragment"
misc_feature      70415..77680
                  /note="assembly_fragment"

Query Match      17.4%; Score 412.6; DB 2; Length 129261;
Best Local Similarity 95.7%; Pred. No. 5.2e-78;
Matches 424; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy      26      aacggccgcaagtgctgagaaagtgaggagataagaaagcgtggcgggagggcctag 85
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 122477  AATGGCTGCGTGAATCACTGAGAGTGCGGGGATAGAGACGTGCGGGGGCTAG 122536
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      86      cctcgtgagggtccttaagtagcgctcgctgctccctggacagctagtcctaacgac 145
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 122537  CCTCGTGGGGGCTCTTAAAGTAGCGGGCTGCTTCCCTGGCAGCGTACTTAAAGAC 122596
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      146      gtcaagtcagtggtgtcagagctgagcgaagcagctgcagcggccctcttgaa 205
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 122597  GTCAAGGTCAAGTGTGCGAGACTGTGAGAGCAAGCTGAGCGCGCCCTTCTTGAA 122656
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      206      caatgtcccaaccaaggagagctcccaagcctctacgttggtcgttgctagtcggcgaa 265
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 122657  CAATGTCCACACAGAGAGAGCTCCCAAGGCTCTTACGTGGGCTGTGCTAGTCGGCAA 122716
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      266      gctcgaagtgctccgagcgcagagagagcggcgcgagagtgtaacgaagtc 325
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 122717  GCTCGACAGTCTCCGAGCGCCAGAGAGAGCGCGGAGAGAGTTGAATCGAACAGCTC 122776
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      326      ctgaaggataagtgctggcagaggtgctgctggagagagcactctagggcccgagg 385
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 122777  CTGAGGATATATGTCGGCAGAGGTGTCTCTGAGAGGACACTTCTAGCGCCCGGAGG 122836
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      386      ccgtggcgcctgtgtcagcagagagcctgcgtgtgttgcttgtaagaagccacttggtg 445
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 122837  CCGTGGCGCTGTGTGACGAGAACTCGCGCTGTCTTCTTGAAAGAGCCACTTGTGTG 122896
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      446      gcgcgtaacatcggtcgtgtggtg 468
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 122897  GCgcgTAATGCTGTGAGAAATGGG 122919
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

RESULT  11
AC019205/c      198052 bp      DNA      linear      PRI 09-JAN-2002
LOCUS      AC019205
DEFINITION      Homo sapiens BAC clone RP11-398K22 from 6, complete sequence.
ACCESSION      AC019205
VERSION      AC019205.4 GI:12704683
KEYWORDS      HTG.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 198052)
              Sulston, J.E. and Waterston, R.
              Toward a complete human genome sequence
              Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE      99063792
REFERENCE      2 (bases 1 to 198052)
              Nguyen, C., Drone, K., Laplant, Y. and Baum, D.
              The sequence of Homo sapiens BAC clone RP11-398K22
              Unpublished (2001)
REFERENCE      3 (bases 1 to 198052)
              Waterston, R.H.
              Direct Submission
              Submitted (30-DEC-1999) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
              4 (bases 1 to 198052)
              Submitted (07-FEB-2001) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
              5 (bases 1 to 198052)
              Submitted (09-JAN-2002) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
              On Feb 7, 2001 this sequence version replaced gi:17630978.
              ----- Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu/gsc
              Contact: saplens@wustl.wustl.edu
              ----- Summary Statistics
              -----
              Center project name: H_NH0398K22
              -----

```

REFERENCE

MO 63108, USA
5 (bases 1 to 198052)
Waterston, R.H.
Direct Submission
Submitted (08-FEB-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 198052)
Waterston, R.
Direct Submission
Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Feb 7, 2001 this sequence version replaced gi:17630978.

AUTHORS

Waterston, R.H.

TITLE

Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Feb 7, 2001 this sequence version replaced gi:17630978.

COMMENT

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics

Center project name: H_NH0398K22

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPc1-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Calane, J.J. and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pictet de Jong
and coworkers at the Roswell Park Cancer Institute
(<http://bacpac.med.buffalo.edu>)
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is AC068247. Actual start of this
clone is at base position 1 of RP11-398K22; actual end is at base
position 198052 of RP11-398K22.

The sequence from base position 170120 to 170941 is represented by
PCR from clone DNA.

FEATURES

location//Qualifiers
1..198052
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6"
/clone="RP11-398K22"
/clone_lib="RPc1-11"
24..325
/rpt_family="Alu"
repeat_region
560..656
/rpt_family="Alu"
repeat_region
665..972
/rpt_family="Alu"
repeat_region
992..1156

repeat_region /rpt_family="Alu" 1163.1388
repeat_region /rpt_family="L2" 1389.1682
repeat_region /rpt_family="Alu" 1683.1756
repeat_region /rpt_family="L2" 1931.2408
misc_feature note="similar to EST R08522 (NID:9760445) ye95b11.r1"
repeat_region /rpt_family="Alu" 2301.2493
repeat_region /rpt_family="Alu" 2630.2940
repeat_region /rpt_family="Alu" 3610.3721
repeat_region /rpt_family="MIR" 4050.4292
repeat_region /rpt_family="Alu" 4846.5144
repeat_region /rpt_family="Alu" 5463.5767
repeat_region /rpt_family="Alu" 5959.6126
repeat_region /rpt_family="Alu" 6208.6327
repeat_region /rpt_family="L1" 6381.6691
repeat_region /rpt_family="Alu" 6752.7035
repeat_region /rpt_family="Alu" 7216.7343
repeat_region /rpt_family="L1" 7347.7428
repeat_region /rpt_family="Alu" 7450.7527
repeat_region /rpt_family="Alu" 7767.7902
repeat_region /rpt_family="Alu" 8078.8284
repeat_region /rpt_family="L1" 8292.8383
repeat_region /rpt_family="L1" 8409.8512
repeat_region /rpt_family="Alu" 8515.8614
repeat_region /rpt_family="L1" 8615.8904
repeat_region /rpt_family="Alu" 8905.9203
repeat_region /rpt_family="L1" 9235.9336
repeat_region /rpt_family="MER2_type" 9333.9393
repeat_region /rpt_family="MER2_type" 9397.9661
repeat_region /rpt_family="L1" 9662.9774
repeat_region /rpt_family="Alu" 9775.10076
repeat_region /rpt_family="Alu" 10077.10088
repeat_region /rpt_family="Alu" 10105.10297
repeat_region /rpt_family="Alu" 10302.10450
repeat_region /rpt_family="L1" 10458.10758
repeat_region /rpt_family="Alu" 10494.10753
misc_feature note="similar to EST BE890623 (NID:910349131)"
repeat_region /rpt_family="Alu" 11086.11092
repeat_region /rpt_family="Alu" 11101.11407
repeat_region /rpt_family="Alu"

repeat_region 11439.11725 /rpt_family="Alu"
repeat_region 11728.11916 /rpt_family="L1"
repeat_region 11949.12059 /rpt_family="L1"
repeat_region 12060.12371 /rpt_family="L1"
repeat_region 12372.12406 /rpt_family="Alu"
repeat_region 12431.12740 /rpt_family="L1"
repeat_region 12745.12898 /rpt_family="Alu"
repeat_region 12899.13211 /rpt_family="L1"
repeat_region 13212.13612 /rpt_family="Alu"
repeat_region 13724.14014 /rpt_family="L1"
repeat_region 14341.14655 /rpt_family="Alu"
repeat_region 15327.15489 /rpt_family="Alu"
repeat_region 15534.15738 /rpt_family="Alu"

Query Match 17.4% Score 411. DB 9: Length 198052;
Best Local Similarity 95.5% Pred. No. 1.2e-77;

Matches 423; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 26 aacggccgcagtcgtgctgtaaaagatgagcggatagagacgtgagcgggggctag 85
DB 48837 AATGCGCTCGCTGAATCAATCGCAATGCGCGGATAGAGCGGCGGGGCTAG 48778
QY 86 cctggcggggtccttaagtagcggcgtggtgcttcctggaagccttagcttaagac 145
DB 48777 CCTGTCGCGGCTCTTAAGTAGCGGCTGCTGCTTCCCTGCGAGCTACTTACGAC 48718
QY 146 gtccaggtcaggtggtgcagagctlgagcgggaacgacgttcgacgcgcctcttgaa 205
DB 48717 GTCAAGCTCAGTGTGCTGAGACGTGAGCGCAAGCAAGTGCAGCCCTCTTGAA 48658
QY 206 caatgctccacacagagagagctcccaagcctctcgtgggtgttctatgtcggaa 265
DB 48657 CAATGCTCCACACAGAGAGAGCTCCCAAGGCTCTACGTGGTCTTGATGCGGAA 48598
QY 266 gctcgacagtcgtcccgagcgccagagagagagccggcggaggagtgaatcgaaacagtc 325
DB 48597 GCTCGACAGTGTCCCGAGCGCCAGAGAGAGCGCCGCGAGAGTGAATGAAACAGTTC 48538
QY 326 ctggagatatagttgctgcagaggtgctgctgagagagcaacctataggcccgagg 385
DB 48537 CTGAGGATATATAGTTCGCGAGAGAGTGTCTCTGAGAGGACACTTAGGCCCGGAGG 48478
QY 386 ccgtggcgcgtgtgtaacagaaactgcgcgtgtgttctcttgaaagcaacttgctg 445
DB 48477 CCGGCGCGCTGTCTACGAGAACTCCGCTGTCTTGTGAAGACCACTTGTGTC 48418
QY 446 ggcgggaatcgctgctgtgtgg 468
DB 48417 GCGCGATATCGTGTGAGATGCG 48395

RESULT 12
AC010091/c 56864 bp DNA linear PRI 30-SEP-2000
LOCUS Homo sapiens BAC clone RP11-295A1 from 7q11.23-q21.1, complete
DEFINITION
ACCESSION AC010091
VERSION AC010091.2 GI:9211377
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1	(bases 1 to 56664)
TITLE		Toward a complete human genome sequence
JOURNAL		Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE	99063792	
REFERENCE	2	(bases 1 to 56664)
AUTHORS		Cordum,H., Stoneking,T. and Glaser,E.
TITLE		The sequence of Homo sapiens BAC clone RP11-295A1
JOURNAL		Unpublished
REFERENCE	3	(bases 1 to 56664)
AUTHORS		Waterston,R.H.
TITLE		Direct Submission
JOURNAL		Submitted (11-SEP-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis MO 63108, USA
REFERENCE	4	(bases 1 to 56664)
AUTHORS		Waterston,R.H.
TITLE		Direct Submission
JOURNAL		Submitted (15-JUL-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis MO 63108, USA
REFERENCE	5	(bases 1 to 56664)
AUTHORS		Waterston,R.
TITLE		Direct Submission
JOURNAL		Submitted (30-SEP-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT		On Jul 15, 2000 this sequence version replaced g1:15870309.

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GTE/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPI1-*all-human* BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tatenno, M., Catanesse, J. J. and de Jong, P. J. (1998). An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>).

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RPL1-5B9, 200 bp overlap; the

FEATURES
source
Location/Qualifiers
I. .36864

UNRES	Location/Qualifiers
source	1. .56864
	/organism="Homo sapiens"
	/db_xref="taxon:.9606"
	/chromosome="7"
	/map="q11.23-q21.1"
	/clone="RP11-2953A1"
	/clone_id="RPC1-11"
	3. .631
repeat_region	/rpt_family="MERL_type"
	25. .38
misc_feature	/note="similar to EST AI241674 (NID:g3637071) qu70a11.x1"
	25. .38
misc_feature	/note="similar to EST AA716347 (NID:g2728621) zh29a04.sl"
	25. .38
repeat_region	/rpt_family="MERL_type"
	662. .879
repeat_region	/rpt_family="MERL_type"
	883. .1241
repeat_region	/rpt_family="L2"
	2702. .2958
repeat_region	/rpt_family="MER2_type"
	2959. .3075
repeat_region	/rpt_family="Alu"
	3076. .3128
repeat_region	/rpt_family="(CA)n"
	3129. .3297
repeat_region	/rpt_family="Alu"
	3298. .3365
repeat_region	/rpt_family="MER2_type"
	3511. .3808
repeat_region	/rpt_family="Alu"
	4452. .4841
repeat_region	/rpt_family="MALR"
	5091. .5183
repeat_region	/rpt_family="MALR"
	5389. .5530
repeat_region	/rpt_family="L2"
	5849. .6295
repeat_region	/rpt_family="MALR"
	6611. .8631
repeat_region	/rpt_family="L2"
	8859. .8883
repeat_region	/rpt_family="AT_rich"
	9579. .9668
repeat_region	/rpt_family="L2"
	10294. .10473
misc_feature	/note="similar to EST AI582192 (NID:g4568089) tq63a11.x1"
	10379. .10473
misc_feature	/note="similar to EST AI871836 (NID:g5545885) wm51h12.x1"
	10518. .10896
misc_feature	/note="similar to EST AI582192 (NID:g4568089) tq63a11.x1"
	10518. .10901
misc_feature	/note="similar to EST AI871836 (NID:g5545885) wm51h12.x1"
	10539. .10896
misc_feature	/note="similar to EST AI189368 (NID:g3740577) qd05a04.x1"
	10644. .10891
misc_feature	/note="similar to EST AI241674 (NID:g3637071) qu70a11.x1"
	10644. .10897
misc_feature	/note="similar to EST AA716347 (NID:g2728621) zh29a04.sl"
	11201. .11585
repeat_region	/rpt_family="L1"
	12448. .12535
repeat_region	/rpt_family="AT_rich"
	13589. .13730
repeat_region	/rpt_family="MERL_type"
	13762. .13805
repeat_region	/rpt_family="(TTATA)n"
	13940. .13979
repeat_region	/rpt_family="AT_rich"
	14039. .14859
repeat_region	/rpt_family="MERL_type"

```

repeat_region      15600..16114
                    /rpt_family="MER1_type"
repeat_region      16171..16901
                    /rpt_family="MER1_type"
repeat_region      17358..17955
                    /rpt_family="MER2_type"
repeat_region      18111..18340
                    /rpt_family="MER2_type"
repeat_region      18487..18515
                    /rpt_family="AT-rich"
repeat_region      18628..18965
                    /rpt_family="Retroviral"
repeat_region      19321..19352
                    /rpt_family="MIR"
repeat_region      19815..19974
                    /rpt_family="L2"
repeat_region      20255..20319
                    /rpt_family="L2"
repeat_region      20909..22180
                    /rpt_family="Mariner"
repeat_region      22181..22200
                    /rpt_family="TG)n"
repeat_region      22211..22395
                    /rpt_family="Mariner"
repeat_region      23854..23931
                    /rpt_family="AT-rich"
repeat_region      23556..25630
                    /rpt_family="MER94"
repeat_region      26415..26455
                    /rpt_family="TGAA)n"
repeat_region      26562..26843
                    /rpt_family="Alu"
repeat_region      26847..26900
                    /rpt_family="TAA)n"
repeat_region      26989..27117
                    /rpt_family="MER1_type"
repeat_region      27433..27561
                    /rpt_family="MIR"
repeat_region      28264..28530
                    /rpt_family="MIR"
repeat_region      31086..31174
                    /rpt_family="AT-rich"
repeat_region      32846..33080
                    /rpt_family="MIR"
repeat_region      33877..33948
                    /rpt_family="MALR"
repeat_region      /rpt_family="MALR"

Query Match      15.5%; Score 366.8; DB 9; Length 56864;
Best Local Similarity 84.8%; Pred. No. 3.3e-68;
Matches 448; Conservative 0; Mismatches 72; Indels 8; Gaps 3;

QY 126 tggcagctagcttctaacgacgtcagtcagtggtgtagagagctgagcgcaacgacgtc 185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 15575 TCGCACCTAGCTAGTACAGCGACCGTCAGATGCTGCGAGCGCGGGAACGAACTC 15516
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 186 ggaagcgccctctcttggaacaatgctccacaacgagagagctcccaagcctcaagt 245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 15515 GGACGCGCTATTTTGGAAACAAAGTCTCTCCAMGAGAAACCTCCAGCCTCTACATG 15456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 246 gttcgttagctagtcgcgcgaagctcgacagtgctcccgagcgcgacgagagagcgga 305
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 15455 GGTCTGTGCTAGCTGCTGAGAGCTCGACAGTGTCCGGGCAACCGAGAGCGCGCGGA 15396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 306 ggaattgaatcgacaacagtgctctgaaggatatagtgtcgagaaaggttgt---cgctgag 362
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 15395 GGAAGTGAAATCGAAGAGGCTCTGAGGACATAGGACAGCGCTGAGGTGTGGCGCTGGAG 15336
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 363 aggaaccttagagcccccggagcg- tggccgctggttcaagagaactgacgtctgtct 421
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 15335 AGGCCCTATAGGCTCCAGAGGCGGTGGCTGCTGCTACCGGAACTGCCGCAATGTT 15276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 422 ttgcttgaagagccacttgtgtgagcggttaacggttgt----gttggttcaaaata 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 15275 TTGCTTTGAGAACCACTTTGTGCTGTGATGGGTCCGCGCGGTGGTCAAAATA 15216
QY 478 aagaataacaaagtacacaaaccccaatcccaataataacaaacacgaatca 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 15215 AAGATATATACAAAGTACAAACACCCCAATCTTAAACATATACAGAACATATGAAATCA 15156
QY 538 ttatgcaaaacttttggcagcaaggcaatgcgaacgaagaagaacagtcagtgacacat 597
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 15155 TTACTCAAAATTTTAGCCGCAAGCATGTGCAAAACGCAAGCGGATGATGATAT 15096
QY 598 ttctgaaaaagctagaagaataattacattcagaatgcggaattgat 645
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 15095 TTTATTTAAAGCCAGCAAGAAACTACATTTGCAATGCGAATGCGACTTGAT 15048

```

```

RESULT 13
AY051831
LOCUS      2370 bp mRNA linear INV 27-AUG-2001
DEFINITION Drosophila melanogaster LD33749 full length cDNA.
ACCESSION AY051831
VERSION AY051831.1 GI:15291972
KEYWORDS  FLYCDNA.
SOURCE      fruit fly.
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 2370)
            Stapleton,M., Brockslein,P., Hong,L., Agbayan,I.A., Carlson,J.,
            Champe,M., Chavez,C., Dorsett,Y., Farfan,D., Frise,E., George,R.,
            Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,
            Nuno,J., Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K.,
            Yu,C., Lewis,S.E., Rubin,G.M. and Celisner,S.
            Direct Submission
            Submitted (10-AUG-2001) Berkeley Drosophila Genome Project,
            Lawrence Berkeley National Laboratory, One Cyclotron Road,
            Berkeley, CA 94720, USA

```

```

COMMENT
TITLE      JOURNAL
JOURNAL     Submitted (10-AUG-2001) Berkeley Drosophila Genome Project,
            Lawrence Berkeley National Laboratory, One Cyclotron Road,
            Berkeley, CA 94720, USA
            Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory
            Berkeley, CA 94720

```

This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our web site (<http://fruitfly.berkeley.edu>) or send email to conae@fruitfly.berkeley.edu.

FEATURES
source
Location/Qualifiers
1..2370

gene
/organism="Drosophila melanogaster"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"
/map="84E12-84E12"
1..2370

CDS
/gene="CG7878"
/note="Alignment with genomic scaffold AE003677"
/db_xref="FLYBASE:FBgn0037549"
99..2210

/gene="CG7878"
/note="Longest ORF"
/codon_start=1
/db_xref="FLYBASE:FBgn0037549"
/product="LD33749p"
/protein_id="AAK93255.1"

```

      /db_xref="GI:1529197.3"
      /translation="MSLDMDDDPNVEYKAVQERYSKPNKYEKRTSRDGNRAKGGD
      RNTDGGEGNGTFAKKNIDYGGRRDNDNTGACHRSDRSGGGGGGSPSSUSIAEMK
      GKATGGGNSINRITONDPNVKFLDKCDLIVATITSIQSNVITAIVHVKQVTSNGD
      GQDRDRHDGSHNERTDGGDYDRYKAGNSYFENNTSESKNGDGLTETIDWELINKAT
      AATAWMSKSCPLTLNFKYKAPVANULTKSELRREBNKTHYVSVFPEKGEYSPP
      IIPWTVPEQCAFEPYKLEIKETIKAVNLSKSPISQOAMPILLOGMDSKTLTGVM
      LALDLGMIHTEYQSTPRGRCGANLVLAIPRELAQIDIEBKYSFPMGNAMCVVGG
      GGRNNQIDLERGAEIITCTPGRLANDLVIAVLDVITVLVADABDMLMGEEEP
      IRVYMLDIPRODTMTSATMPGVRRLLASVKMPKPIOVGCSIDLAATHSKQITK
      MEDDMKFWTITSPFKYMNSSDKIIITFGKRYKRDLSSETLTDGNGCTIGNNDG
      DPEOALADIKSCVYVILVATDVASRGDLIEDITHYVNDPPHNIEYVHVRGCRG
      RQGTSSFTPRRDNMAKLELITLOABDQVYDELHMNRRFAMNDKALAEGGGFGGG
      RGRFEGGGGGRRNDFDFQI"
BASE COUNT      669 a      534 c      634 g      533 t
ORIGIN

```

Query Match	15.18;	Score 358;	DB 3;	Length 2370;
Best Local Similarity	55.58;	Pred. No. 2.2e-66;		
Matches 741;	Conservative 0;	Mismatches 580;	Indels 15;	Gaps 2;
Qy 764 caaagtgaggagacttaccacacatbaagaaacatttlaaagagtcacgtccacaa	823			
Db 766 CAGGGGAGCCAAAGTCCACCGCTTAACAAGAACTTTTACGAAGGCGCTCGAATGG	825			
Qy 824 gtcgcattcaaaagtgaagcagatagttgaggaagaagaatttaataaac-----	878			
Db 826 CCANATGTGCAAAATCGAGATTGACCGCATTTCCGAGGAGAAACAAGATTAACACTTT	885			
Qy 879 -----gtggagtagactgaagatbvgggagaacgaacctatcccccaatccctacgca	931			
Db 886 CTATATTTTTGAGACCCAAAGAGGGGGAAGCTGACCGCTATTCCAAATCCAGTTTGA	945			
Qy 932 catttgaagaagcgtcttcaatgtaatacctggaggtatagaacaatataaaggcaggt	991			
Db 946 CCTTGAGCAATGCTTTTGGCGAATATCCAGATATGCTTTGGAGAGATTACAAATGGGCT	1005			
Qy 992 ttcaaaagccaacacctatagtcacagcagtcgccaattgltgcaaggaaatagac	1051			
Db 1006 TCTCCAGAGCCCTCCCATCATTCAGAGCGGTGGCCCATTTCTCCAGAGTCATACGA	1065			
Qy 1052 ttaagaagtagtcagcagacttgtaacagaagaacatgvtgttaattaaagcttgattta	1111			
Db 1066 TGAATCGATATCGCCCAACAGCAGCAGTGGCAAAACATAGCTTCTTTTCCCGGATGA	1125			
Qy 1112 ttcatctgctcccttcaacccagcctttaaagvtcaaggaaatagaaccgycgatbgtatc	1171			
Db 1126 TTTACACGAGTACACAGCAGCAGCCAGAGGACAGAGGGGCGAGGTAACTGTTGGTGC	1185			
Qy 1172 taaactccacactccggaaatlagacacttcaagttaagaagaatgttgcaaatcatata	1231			
Db 1186 TGGCAGCCACTCGGGAATTTGGCTCTGCAAAATCGAAGTGAAGGTATCTCTTTC	1245			
Qy 1232 aagggtctcggaaigtgttgtatagatbgtgvtgaaatagaagaatgaacaatagaagac	1291			
Db 1246 GTTGCGATGTGAAGCGTTTGGTGTACGGGGGTGGCATTCGCAACATGACGATATCCACT	1305			
Qy 1292 tttaaaaagaagttagatatacatataatgtgaactcccggaagaatgtgaatgactgcgaatga	1351			
Db 1306 TGGAGCGCGGTGCGAATCATATGTGACTCCGGGAGCCCTGAAACATTTGATTAATGG	1365			
Qy 1352 gtaacttcgtcaactcgcagaatatabaactacttgttltttagatgaagcagaagaatgt	1411			
Db 1366 CCANATGCTTGAATGTAAAGCAACCATACTATTACTGTGGTTCGAGAGAGATATGCGATGC	1425			
Qy 1412 tgaacaatggaatttgaaccccaataatgaatatttgcattagaatgvggcccagaatagc	1471			
Db 1426 TGGACATATGGGTTTTCGAGCCGCAATTCGGAAGGTGATGCTAAGCATACCTCGGAGCGTC	1485			
Qy 1472 agacagtatagcaagtgctacatggtcctcatcaatgtcatcgtccgcacaaatcatt	1531			
Db 1486 AGACCATATATGAGCTGGCCACTTTGGCCACAGAGATACCTCTTTAAGGACAGAGCATATA	1545			

Qy	1532	tgaaagaccacatgatctgtctatggtggaacttggcctcagttgtctgtgaagttcaatga	1591
Db	1546	tgaaagatcccatccacagctgtgtgtggaatgccttcgatctggcagccacacacactgcgtga	1605
Qy	1592	agcaaaataaattctgtaaccaccgga-ga--agagaaatgagtcacatgcaaaccttcc	1648
Db	1606	AAACAATTTTAAATTTGATGTGAGAGGATGACATGTGACAAATTCACACACATTTACATCTTTCG	1665
Qy	1649	tacagagatctgtcaccacagacaagaatcattgtcttcgttctctcgaaaagctgttcgag	1708
Db	1666	TTTAAGAACATGTCCAGTACGACACAAATGATCATATTTTGTGGAGCACAAGTGTGTCTGT	1725
Qy	1709	atcaactatcaagatgacctaaactctggaatatatacgaagaagctctctgcattggagata	1768
Db	1726	ACGACCTATACCGATGACCTTTACGCTGTGAGTGTTCATGTACCCAGTGCATTCATGTGATATC	1785
Qy	1769	gagaaacagagatcggagagaagacatacgaagaacttaaaacaaaggaaggaatgaac	1828
Db	1786	GGCATCAGATGTGATGTGAGACAGAGCTATTTGCCGATATTAAATGTCGGCGTGGCGCATTC	1845
Qy	1829	taattgcaactgatactgctctcagaagagcttgatgftccatgacgttacaagtctacata	1888
Db	1846	TGGTTCCTTACCGATGTGTGACATCACTGCTGTGACATTTGAGATATGCACATCTGTCATCA	1905
Qy	1889	attctgacttccacggaatatctggaagaatacgtatacaccgaagaagggcgcagaggaagag	1948
Db	1906	ACTATATATTTTCCGACACACATGTGAGAGATATATGTCTCACCGTGTGTGCTGCACCGAGTGTG	1965
Qy	1949	cagagagagactggtgttccattacaactcttgactaaagaatgttgagaggtgtcctgt	2008
Db	1966	CTGGCCGACAGCGCATCATTAATTAAGCTTTTACGCGCGAGAGATTTGGCTATNGGCAAGG	2025
Qy	2009	aattgataatactctggaagaagcaatcagagatctccaagagagcttgatacatg	2068
Db	2026	AACTAATTTAGATATCTGTGACAGAGCGCGACAGACGAGAGATGTCCCGACGAACATGCACAACTATGG	2085
Qy	2069	ctgagaggtctgaagc 2084	
Db	2086	CTAGACGCTTTAAAGC 2101	
RESULT 14			
AR157511			
LOCUS	AR157511	333 bp	DNA
DEFINITION	Sequence 1 from patent US 6245525.		linear
ACCESSION	AR157511		
VERSION	AR157511.1	GI:16218456	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
ADDITIONAL	1 (bases 1 to 323)		
TITLE	Martelange, V., De Smet, C. and Boon-Fallieur, F.		
JOURNAL	Tumor associated nucleic acids and uses therefor		
FEATURES	Patent: US 6245525-A 1 12-JUN-2001;		
source	Location/Qualifiers		
	1..323		
	/organism="unknown"		
BASE COUNT	120 a	55 c	72 g
ORIGIN			76 t
Query Match	13.0%;	Score 307.8;	DB 6;
Best Local Similarity	98.8%;	Pred. NO. 1.1e-55;	Length 323;
Matches 321;	Conservative 0;	Mismatches 2;	Indels 2;
		Gaps 1.	
Qy	727	gattcaataaaggaaggtttggaatgcaataaagaagaagtcggcagattaccaca	786
Db	1	GATCAATTTAGAGAGAAAGCTTTGAATGTGCAAAAACAAAGTGGCAGATTTTACACACA	60
Qy	787	attaagaaaaatttataaagaagctccactgcacaaagtgccatgctcaaaagtgaaga	846

Db 61 ATTAGAAAACTTTTATTAAGAGTCACAGTCCACAGTGCATGTCAAAAGTAGAAGCA 120
Qy 847 gataagctggagaaagaanaatttaataacgttggaatcttgaaagatgggagaa 906
|||||
Db 121 GATAGTGGAGGAAAGAAATTTTATATACGTGGATGACTGAAGATGGGAGAA 180
Qy 907 ggaactatcccaatcctacatgacatcttgatgaagcccttcaatgttactctgaagt 966
|||||
Db 181 CGACCTATCC--AATCTACCTGCACATTTGATGAGCGCTTCAATGTATCCTGAGGTT 238
Qy 967 atggaacacatlaaagaagcaggcttcaaaagccaacactatcagtcacagcagatg 1026
|||||
Db 239 ATGGAACCATTAAGGCGAGTTTCAAAAGCCACACCTATTCAAGTCACAGGATGG 298
Qy 1027 cccattgtgtgcaaggaatagatc 1051
|||||
Db 299 CCCATTGTGTGCAAGGAAATAGATC 323

RESULT 15

ARL73200

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

120 a 55 c 72 g 76 t

13.08; Score 307.8; DB 6; Length 323;

Best Local Similarity 98.8%; Pred. No. 1.1e-55;

Matches 321; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Query Match

Db 727

1

GATCAATTAAGAGGAGGTTGAAATGCAAAACAAAGTGGCAGATTACCA 60

787

attagaaacatttataaagagtcacatgcccacagtgccatgltcaaaatagaagca 846

61

ATTAGAAAACTTTTATTAAGAGTCCACTGCACAGTGCATGTCAAAAGTAGAAGCA 120

847

gataagctggagaaagaanaatttaataacgttggaatcttgaaagatgggagaa 906

121

GATAGTGGAGGAAAGAAATTTTATATACGTGGATGACTGAAGATGGGAGAA 180

907

cgaactatcccaatcctacatgacatcttgatgaagcccttcaatgttactctgaagt 966

181

CGACCTATCC--AATCTACCTGCACATTTGATGAGCGCTTCAATGTATCCTGAGGTT 238

967

atggaacacatlaaagaagcaggcttcaaaagccaacactatcagtcacagcagatg 1026

239

ATGGAACCATTAAGGCGAGTTTCAAAAGCCACACCTATTCAAGTCACAGGATGG 298

1027

cccattgtgtgcaaggaatagatc 1051

299

CCCATTGTGTGCAAGGAAATAGATC 323

Search completed: July 25, 2002, 12:29:30
Job time: 11910 sec

•
•

•

•